



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 162984

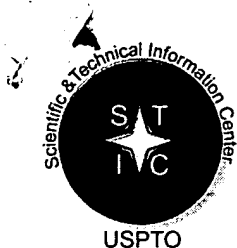
TO: Phuong Bui
Location: REM-2A15&2C18
Art Unit: 1638
Friday, September 02, 2005
Case Serial Number: 10/762049

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

barbara.obryen@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 165364

TO: Phuong Bui
Location: rem/2A15/2C18
Art Unit: 1638

Sept 14, 2005

Case Serial Number: *10/762049*
~~10/62049~~

From: P. Sheppard
Location: Remsen Building
Phone: (571) 272-2529

sheppard@uspto.gov

Search Notes

Interference

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 23:27:41 ; Search time 10582 Seconds
(without alignments)
11214.026 Million cell updates/sec

Title:	US-10-762-049-17
Perfect score:	2449
Sequence:	1 gcacgagctagctcgacat.....aaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Minimum	DB seq	length:	0
Maximum	DB seq	length:	2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum match 100%
Listing first 45 summaries

Database :

```
Database :
Genre:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	2449	100.0	2449	6	AR477464	AR477464 Sequence	
2	1042.8	42.6	2045	8	D89631	D89631 Arabidopsis	
3	1041.2	42.5	2266	8	AY099873	AY099873 Arabidopsis	
4	1032	42.1	2193	6	BN4581745	BN4581745 Brassica	
5	1018.2	41.6	1977	6	AX506807	AX506807 Sequence	
6	1018.2	41.6	1977	6	AX851748	AX851748 Sequence	
7	897.4	36.6	2037	8	AB004060	AB004060 Arabidopsis	
8	864.8	35.3	2219	8	AF601439	AF601439 Brassica	
9	846	34.5	1974	6	AX653455	AX653455 Sequence	
10	846	34.5	2327	8	AK066932	AK066932 Oryza sat	
11	765.2	31.2	1981	6	AR477463	AR477463 Sequence	
12	720.6	29.4	2067	6	AR477463	AR477463 Sequence	
13	681	27.8	2259	8	AJ704374	AJ704374 Brassica	
14	678.6	27.7	2256	8	AB054645	AB054645 Arabidopsis	
15	677.4	27.7	2640	8	AF355602	AF355602 Zea mays	
16	667.8	27.3	2181	8	AK104831	AK104831 Oryza sat	
17	666.2	27.2	2338	8	AK072809	AK072809 Oryza sat	
18	660.2	27.0	1896	6	AX651700	AX651700 Sequence	
19	659.2	26.9	1917	8	AB023423	AB023423 Arabidopsis	

20	559.2	26.9	2470	8	A3347614	Lycopodium
21	555.4	26.8	2279	6	AR477456	Sequence
22	553.4	26.7	2159	8	AB042322	Arabidopsis
23	550.4	26.6	2439	8	AF309643	Solanum
24	550.4	26.6	1989	6	AX653919	Sequence
25	650.4	26.6	2791	8	AK103007	Oryza sativa
26	649.8	26.5	2262	8	BN416460	Brassica
27	647.2	26.4	1947	6	AX653223	Sequence
28	647.2	26.4	1989	8	AF493792	Oryza sativa
29	646	26.4	2167	8	BN413388	Brassica
30	640.4	26.1	2310	8	BT009249	Triticum
31	640.4	26.1	2311	6	AR477465	Sequence
32	639	26.1	2248	8	AF332959	Lycopodium
33	637.4	26.0	2327	8	AF347613	Lycopodium
34	634.4	25.9	1971	6	BD175021	Gene encoding
35	634.4	25.9	1971	8	AB049624	Arabidopsis
36	632.6	25.8	2494	8	AK067270	Oryza sativa
37	630.4	25.7	1940	6	AX654768	Sequence
38	627	25.6	2442	8	HVHVST1	Hordium vulgare
39	622	25.4	2137	8	SHSHST2	S. hamata SH
40	620.6	25.3	2456	8	HVH52867	S. hamata SH
41	614	25.1	2454	8	SHSHST1	S. hamata SH
42	610	24.9	2515	8	SSSULPTRP	Staphylinus
43	609.8	24.9	2313	8	AB018695	Arabidopsis
44	600.2	24.5	2000	6	AX653563	Sequence
45	586.4	23.9	2808	8	AK121195	Oryza sativa

ALIGNMENTS

RESULT	1
LOCUS	AR477464
DEFINITION	Sequence 17 from patent US 6696292.
ACCESSION	AR477464
VERSION	AR477464.1 GI:47234922
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified. 1 (bases 1 to 2449)
AUTHORS	Allen,S.M., Falco,S.C. and Thorpe,C.J.
TITLE	Genes encoding sulfate assimilation proteins
JOURNAL	Patent: US 6696292-A 17 24-FEB-2004;
FEATURES	Location/Qualifiers 1..2449 /oranism="genomic" /mol_tvsos="DNA"
source	

ORIGIN		Query Match	100.0%;	Score 2449;	DB 6;	Length 2449;
		Best Local Similarity	100.0%;	Pred. No. 0;		
		Matches 2449;	Conservative	0;	Mismatches	0; Indels
						0; Gaps
Qy	1	GCACGAGCTAGCTCGCACATTAAAGTTATATAACACATATTTGCTTGCTTGAATACTAT	60			
Db	1	GCACGAGCTAGCTCGCACATTAAAGTTATATAACACATATTTGCTTGCTTGAATACTAT	60			
Qy	61	TATTGAAGATATGCGGAGTGTAGATTATGATGATACCCCTTTGGGCATGAACAACCTTTTGAGAG	120			
Db	61	TATTGAAGATATGCGGAGTGTAGATTATGATGATACCCCTTTGGGCATGAACAACCTTTTGAGAG	120			
Qy	121	AGTGCAACCAAGTCGAGGTTCCACGGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTT	180			
Db	121	AGTGCAACCAAGTCGAGGTTCCACGGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTT	180			
Qy	181	GAAGGAGACTTTCTTCCCTGATGACCTTTTGAGGCAGTTCAAGAAACCAAGCCAGCTTCCAA	240			
Db	181	GAAGGAGACTTTCTTCCCTGATGACCTTTTGAGGCAGTTCAAGAAACCAAGCCAGCTTCCAA	240			
Qy	241	GAAGTTTCATGCTTGGGCTTCAGTCTCTTCTTCCCATTTTCGAATGGGCTCCCAATAACAC	300			

Db 241 GAAGTTCATGCTTGGCGCTTCAGTTCCTCTTCCCAATTTTCGAATGGGCTCCAAAATACAC 300
Qy 301 CTTTCAGTTCCTTAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTC 360
Db 301 CTTTCAGTTCCTTAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTC 360
Qy 361 TCAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTTCGGACTATATTCGAG 420
Db 361 TCAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTTCGGACTATATTCGAG 420
Qy 421 CTTTATACCAACCATGATTTATCGGATGATGGTAGCTCGAGGATTTGGCAGTGGGAC 480
Db 421 CTTTATACCAACCATGATTTATCGGATGATGGTAGCTCGAGGATTTGGCAGTGGGAC 480
Qy 481 TGTGGCGGTTGGATCGCTTCGATGGGTTTCGATGTTGAGTAATGCCGTTGATCCCAATGA 540
Db 481 TGTGGCGGTTGGATCGCTTCGATGGGTTTCGATGTTGAGTAATGCCGTTGATCCCAATGA 540
Qy 541 AGACCCAAAGCTTTACCTCCACTGGCTTTTCACAGCTACATTAATTTGCTGGTGTTC 600
Db 541 AGACCCAAAGCTTTACCTCCACTGGCTTTTCACAGCTACATTAATTTGCTGGTGTTC 600
Qy 601 GGCCTGCTGGGTCCTGTTAGTGGGTTGATCGTGGATTTCTGTCACATGCAACCAT 660
Db 601 GGCCTGCTGGGTCCTGTTAGTGGGTTGATCGTGGATTTCTGTCACATGCAACCAT 660
Qy 661 AATAGGGTTTCATGGGAGGACGACCGGTGGTGTCTGCAGCAACTAAATCGAATCT 720
Db 661 AATAGGGTTTCATGGGAGGACGACCGGTGGTGTCTGCAGCAACTAAATCGAATCT 720
Qy 721 TGGCCTTGAGCATTTCAACCATGGAGCTGATATCATATCAGTGATCGCTCTGTTTTCAC 780
Db 721 TGGCCTTGAGCATTTCAACCATGGAGCTGATATCATATCAGTGATCGCTCTGTTTTCAC 780
Qy 781 CCAAACTCATGAGTGGAGTGGAAAGTCTGTGTTAGGATGCTTCATTTCTCTCT 840
Db 781 CCAAACTCATGAGTGGAGTGGAAAGTCTGTGTTAGGATGCTTCATTTCTCTCT 840
Qy 841 CTTAGCACAAAGATCTTCAGCAAAAAACGACCAAGGTTTTTTGGGTGTGAGCAATGCG 900
Db 841 CTTAGCACAAAGATCTTCAGCAAAAAACGACCAAGGTTTTTTGGGTGTGAGCAATGCG 900
Qy 901 GCCATTGACGTCGGTTATATGGGAAAGTCTCTTGGTTATTTCACTACGCGGAGAAGCA 960
Db 901 GCCATTGACGTCGGTTATATGGGAAAGTCTCTTGGTTATTTCACTACGCGGAGAAGCA 960
Qy 961 CGGTGTTGAAGTGATAGGAGAACTGAAGAAGGGTTTGAATCCACCATCACTCACAATCT 1020
Db 961 CGGTGTTGAAGTGATAGGAGAACTGAAGAAGGGTTTGAATCCACCATCACTCACAATCT 1020
Qy 1021 GGTATTGTTGTCGCCCTTACATGACTACAGCTGTCAAAACTGGCATTTGCTGGCATCAT 1080
Db 1021 GGTATTGTTGTCGCCCTTACATGACTACAGCTGTCAAAACTGGCATTTGCTGGCATCAT 1080
Qy 1081 ATCATTTGGGAAGGAATAGCATAGGAAGAAGCTTTTGCATATATAAAAAATTAACAATAT 1140
Db 1081 ATCATTTGGGAAGGAATAGCATAGGAAGAAGCTTTTGCATATATAAAAAATTAACAATAT 1140
Qy 1141 TGATGGCAACAAGAGATGATAGCTATTGGGACCATGAACGTTAGTTGTTCTTACCTC 1200
Db 1141 TGATGGCAACAAGAGATGATAGCTATTGGGACCATGAACGTTAGTTGTTCTTACCTC 1200
Qy 1201 TTGCTACCTCACAAAGGACCAATTTTCGGGTTTCGGCTGTGAACATAAACGCTGGATGCAA 1260
Db 1201 TTGCTACCTCACAAAGGACCAATTTTCGGGTTTCGGCTGTGAACATAAACGCTGGATGCAA 1260
Qy 1261 GACAGAGCTTCCAAATTAATATGATGCTTGGAGTAATGTTGACATTTGTTATCTCTGAC 1320
Db 1261 GACAGAGCTTCCAAATTAATATGATGCTTGGAGTAATGTTGACATTTGTTATCTCTGAC 1320
Qy 1321 ACCCTTGTTCATTTACACTCCCTGGTGGTCTATCAGCTATTATCGTATCTGCAATGCT 1380
Db 1321 ACCCTTGTTCATTTACACTCCCTGGTGGTCTATCAGCTATTATCGTATCTGCAATGCT 1380

Qy 1381 TGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAAATTTGACTTTGT 1440
Db 1381 TGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAAATTTGACTTTGT 1440
Qy 1441 GGTGTGCATGAGTGATGATCAATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAT 1500
Db 1441 GGTGTGCATGAGTGATGATCAATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAT 1500
Qy 1501 AGCTATTGTAATATCTGTACTTTCCGGTACTTCTATTTAATTCGAAGCCCAAGACATTCGT 1560
Db 1501 AGCTATTGTAATATCTGTACTTTCCGGTACTTCTATTTAATTCGAAGCCCAAGACATTCGT 1560
Qy 1561 TTTTGGGCAACATTTCCAAAATTTCTGTATATACCGAAAATGTTGAGCACTATCAAAAATGCAA 1620
Db 1561 TTTTGGGCAACATTTCCAAAATTTCTGTATATACCGAAAATGTTGAGCACTATCAAAAATGCAA 1620
Qy 1621 ACATGTTCTCGAATGCTAAATTTCTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAG 1680
Db 1621 ACATGTTCTCGAATGCTAAATTTCTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAG 1680
Qy 1681 CTATTTAAGAGAAAGGATCACAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTAC 1740
Db 1681 CTATTTAAGAGAAAGGATCACAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTAC 1740
Qy 1741 AGGGGAGACTAGTTTGCAGTATGTTATTAATTCATATGATGCTGTTTGGAAAATTTGATATAC 1800
Db 1741 AGGGGAGACTAGTTTGCAGTATGTTATTAATTCATATGATGCTGTTTGGAAAATTTGATATAC 1800
Qy 1801 AAGTGGAAATAGTATGCTTTGAAAGAGGTGAAGAAGATTACAGAGAGAAGAGCTACAGCT 1860
Db 1801 AAGTGGAAATAGTATGCTTTGAAAGAGGTGAAGAAGATTACAGAGAGAAGAGCTACAGCT 1860
Qy 1861 TGTTTTGGTCAATCTCTGTAAGTGAAGTGAAGAAACTGAACAAATCGAAGTTCCAAAA 1920
Db 1861 TGTTTTGGTCAATCTCTGTAAGTGAAGTGAAGAAACTGAACAAATCGAAGTTCCAAAA 1920
Qy 1921 TCATTTAGGGAAAGAAATGGATCTATCTGACTGTTTGAAGAGCGCGTTGGAGCATGCAACTT 1980
Db 1921 TCATTTAGGGAAAGAAATGGATCTATCTGACTGTTTGAAGAGCGCGTTGGAGCATGCAACTT 1980
Qy 1981 CAATCTAGCTGCAAGCAAAAAACGAAACCCAAAGAAAGATGAACAGAGGGTTGGAACATGT 2040
Db 1981 CAATCTAGCTGCAAGCAAAAAACGAAACCCAAAGAAAGATGAACAGAGGGTTGGAACATGT 2040
Qy 2041 GTGACTGAGTCATATGCCAAAGAGTATTCTAAATAACTCAAAAAGCTTATTCGTTTCTGT 2100
Db 2041 GTGACTGAGTCATATGCCAAAGAGTATTCTAAATAACTCAAAAAGCTTATTCGTTTCTGT 2100
Qy 2101 CTTAGTAATGTTTACCCTACATGATGTGGCATGAGAAATTTCTGAATCAGCCGGAAGAAGT 2160
Db 2101 CTTAGTAATGTTTACCCTACATGATGTGGCATGAGAAATTTCTGAATCAGCCGGAAGAAGT 2160
Qy 2161 TTTAAAGGCATAGGAAAATGAAAGATGCAAGGGTCTCTAAATTTCTCAACTCTGATCTCT 2220
Db 2161 TTTAAAGGCATAGGAAAATGAAAGATGCAAGGGTCTCTAAATTTCTCAACTCTGATCTCT 2220
Qy 2221 TAGTTAGAGAAAATCTCCTATGATAGCTGTTGAAATTAATCTTTAGTATCATGCTTGT 2280
Db 2221 TAGTTAGAGAAAATCTCCTATGATAGCTGTTGAAATTAATCTTTAGTATCATGCTTGT 2280
Qy 2281 ATAATATATTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACACATG 2340
Db 2281 ATAATATATTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACACATG 2340
Qy 2341 TAAAGAGTAAAGAGTGTGTTGCTAGCACTCTCCATATTCAAATTTGAAAGTAATGGC 2400
Db 2341 TAAAGAGTAAAGAGTGTGTTGCTAGCACTCTCCATATTCAAATTTGAAAGTAATGGC 2400
Qy 2401 ATGAGAAATTTAAATAATCTTTTGGAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2449
Db 2401 ATGAGAAATTTAAATAATCTTTTGGAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2449

```

RESULT 2
D89631
LOCUS       D89631             2045 bp      mRNA      linear      PLN 14-FEB-2004
DEFINITION Arabidopsis thaliana mRNA for sulfate transporter, complete cds.
ACCESSION   D89631
VERSION     D89631.1 GI:2285884
KEYWORDS    ATST1; sulfate transporter.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
Yanaguchi, Y., Nakamura, T., Harada, E., Koizumi, N. and Sano, H.
The electronic Plant Gene Register
Plant Physiol. 113 (4), 1463-1465 (1997)
97267155
REFERENCE   2 (bases 1 to 2045)
AUTHORS    Yanaguchi, Y.
TITLE       Direct Submission
JOURNAL     Submitted (29-NOV-1996) Yube Yamaguchi, Nara Institute of Science
and Technology; 8916-5 Takayama, Ikoma, Nara 630-01, Japan
(B-mail:yu-yamae@mailgate.aist-nara.ac.jp, Tel:07437-2-5652,
Fax:07437-2-5659)
FEATURES             Location/Qualifiers
     1..2045
         /organism="Arabidopsis thaliana"
         /mol_type="mRNA"
         /db_xref="taxon:3702"
         /tissue_type="above-ground parts"
         /clone_lib="lambda ZAP II"
         /ecotype="Landsberg erecta"
     1..2045
         /gene="ATST1"
         /note="33..2009"
         /gene="ATST1"
         /codon_start=1
         /product="sulfate transporter"
         /protein_id="BAA21657.1"
         /db_xref="GI:2285885"
         /translation="MGVEDYTFQGAELHRRHHTVEAQQPFLKSLQYSVKETLFF
DDPFRQPNQAGKRVGLVGLPLPIPEWAPRYNLKPFKSDLIAGITIASLAIPQGIS
YAKLNLPPILGLYSSPPLVAVLGSRLDAGTVAVASLLTGAMLSKEVDAEKDP
KLYLHLAFTATFAGVLEASIGIFRLGFIVDPLSHATIVGPMGAATVSLQQLKGI
GLHFTSDIVISVMRSVFQTHWRWESVGLCGFLFLLSTRIYSIKPKPFVWAA
MPLTSDIGLSLIVFTHAERHQVGTIDKGLNPLSGDLITSPYKSTAVTKTGLI
TGIALAEGTAGVRSFAMFNKYNIDGNKEMIAFQMMNIVGSFTCYLTGTFPFSRVA
YNAGCTAMGNIVMAIATMFLPLTFLPHVTPVLVLSAIIISAMGLGIDYQAAILHW
KVKDFDLVCMSAVGVVGVSGVIGLVAVAISTARLLFVSPKTAVKGNIPNSMIY
KNEQVPSRSTVGGILLIEDAIPYFANASYLRRRIIRWDIEBERVKQSGESLQYI
ILDMASVGNIDTSGISNMWSEIKVIDRRALKVLISNPKGEBVVKKLRSKFIQDHLGKE
WMFLTVEAVEACSYMLHTFTKTPASKNEPNNV"
ORIGIN
Query Match       42.68;      Score 1042.8;  DB 8;  Length 2045;
Best Local Similarity 72.4%;  Pred. No. 3.9e-248;
Matches 1366;  Conservative 0;  Mismatches 517;  Indels 3;  Gaps 1;

Qy 125  CACCAAGTCGAGGTTCCACGCCACACGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG 184
Db 90   CACACGGTGAGGCTCCACAACCTCAACCGTCTCTGAAGTCACTTCAGTACTCAGTGAAG 149
Qy 185  GAGACTTTCTCCTGTATGACCTTTTGAGGAGCTTCAAGAACAGCAGCTTCCAGAAG 244
Db 150  GAAACTCTGTTTCAGACGACCCCTTTTAGACAATTTAAGAACCAAAAATGCATCAAGAAA 209
Qy 245  TTCATGCTTGGCCCTTCAGTCTCTCTCCCACTTTTTCGAATGGGCTCCCAAAATACACCTTT 304
Db 210  TTTGTGTTAGCCCTCAAAATACTTCTCCGATTTTCGAATGGGCACACCGCTACATCTC 269
Qy 305  CAGTTCTTGAAGAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAG 364

```


545 CCAAAGCTTTACCTCCACCTGGCTTTTCACAGTACATTAATTTGCTGGTGTGTTTTTCAGGCT 604
561 CTTAAGCTTTACCTTTACCTTGTCTTTCACCGCCACTTTTTCGCGCGGTTCCTGGAAGCC 620
605 GCCTTTGGTCTGTTTAGGTTGGGGTTGATCGTGATTTCTGTGCATGCAACCATTAATA 664
621 TCTCTTGGAATTTTCAGGTTAGGGTTTATAGTGACATTTCTATCGCATGCAACATAGTA 680
665 GGGTTCATGGAGGAGCAGCACCAGTGTGTGTCTGCAGCAACTAAATCGATTTCTTGGC 724
681 GGATTCATGGAGGAGCAGCGAGCGTGTGAGTCTGCAACAGCTTAAGGGTAATTTTCGGA 740
725 CTTGAGCATTTACCCATGAGAGTGAATATCATATCATGATGAGCGCTCTGTTTTCAACCCAA 784
741 CTTAAACATTTTCAGAGACTCTACCGGATGTTATCTCTGTCTGATCGCTCCGTTTCTCCAA 800
785 ACTCATGAGTGGAGTGGGAAGTGTGTGTTAGGATGTGTTCAITTTCTTCTCTCTCTT 844
801 ACTCAGAGTGGAGATGGGAAGTGGGTTCTTGATGTGGTTTCTATTTCTTCTTCTC 860
845 AGCAAGAGATCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGCCA 904
861 TCCACCATATTTTCAGCATCAAGAAACCAAAATTTCTTTGGGTGGCGGATGGCTCCT 920
905 TTGACGTCGGTTATTTAGGGAAGTCTCTTGGTTTATTTTACTCTACGCGCGAGACGGT 964
921 TTGACCTCAGTGATTTCTTGGGAAGTCTCTTGGTTTACTTCTACTCACTGAGAGACATGGT 980
965 GTTCAAGTGATGAGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAAATCTGGTA 1024
981 GTTCAAGTGATGAGGGACCTGAAGAAGGTTTGAATCCACTCTCCGGTCTTGATCTCATC 1040
1025 TTTGTGTGCGCTTACATGACTACAGCTGTCAAACTGGCAATGTGCTGGGCATCATATCA 1084
1041 TTTACTTCCCCTTACATGTCACAGCTGTCAAACTGGCCTCATCACTGGCATCATTTGCT 1100
1085 CTTCCGGAAGAAATAGCAGTAGGAAGAGCTTTTGCATGTATATAAAATTAATATTTGAT 1144
1101 CTCCTGAAGAGTAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTACAACATAGAC 1160
1145 GGCAACAAGAGATGATAGCTATTGGGACCATGAACGTAGTGTGTTCTTTCACCTCTTC 1204
1161 GGGAAACAAGAGATGATAGCTTTTGGAAATGATGAACATCGTTGGTTCTTTCACATCTGT 1220
1205 TACCTCAACAGACCACTTTTCGCTTTCGCTGTGAACATATAACGCTGGATGCAAGACA 1264
1221 TACCTCAACCGGACCACTTTTCAAGGTTCGCACTGAACTACAAACGGGTTCGAAGACC 1280
1265 GCAGCTTCCAACATTAATGATCACTTGCAGTAATGTTGATGTTATTTCTTGACACCC 1324
1281 GCAATGTCCAACATGATGCGGATTCGCGTTATGTTTCACTCTCTTCTTCTTCTCACACCG 1340
1325 TTGTTCCATTAACCTCCCTGGTGTCTATCGATATTATCGTATCGCAATGCTTGA 1384
1341 CTTTTTCACTACACACACTCGTCTCTCTTCCCATCATCATATCCGAATGCTCGGA 1400
1385 CTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTGTGACAAATTTGACTTTTGGTG 1444
1401 CTCATTGACTATCAAGCTGCCATCCATCTCTGGAAGTTGCAAGTTTCGATTTCTCTGTC 1460
1445 TGCATGAGTGATACATTTGGCGTGGTCTTGGCAGTGTGAAATTTGCTTAGTCATAGCT 1504
1461 TGCATGAGCGCTACGTTTGGGGTGGTATTCGGCAGTGTAGAGATTGGACTCGTCGTAGCG 1520
1505 ATTGTATATCTGTTACTTCCGTTACTTCTATTTATTCGAAGGCCCAAGCATTCGTTTGG 1564
1521 GTGGCGATATCTATAGCGAGGTGTTGTTGTGTGCGAGGCAAAAATTCGGGTGAAG 1580
1565 GGCAACATTCCAAATCTCTGTATATACCGAAATGTTGAGCACTATCAAAATGCAAAACAT 1624
1581 GGAACATACCAACAGCATGATCTATAGGAACACTGAGCAGTAGTACCATCATCAGAAC 1640

QY 1625 GTTCTCGGAATGCTAATTTCTAGAGATTGATGACCAATTTTACTTTGCAATGCGAGCTAT 1684
Db 1641 GTTCTCGGTATTTCTCATCTTTGGAGATTGATGCTCCCATCTTACTTTGCTAATGCCAGTTAC 1700
QY 1685 TTAAGAGAAAGGATCACAAGTGTGATTTGATGAAGAAGAAAGAAATTTAAAGCTTACAGGG 1744
Db 1701 TTGCGTGAAGAATCATAGGTGATTTGATGAAGAAGAGAGAGATTAAACAATCAGGA 1760
QY 1745 GAGACTAGTTTGCAGTATGTTAATTTGATATGATGATGCTGTTTGGAAACATTTGATACAAGT 1804
Db 1761 GAGAGCAGCTTACAAATATATTACTCGATATGTCAGTCTTGTGTAATATCGACACAAGC 1820
QY 1805 GGAATAGATGCTTTGAAGAGGTGAAGAAGATTACAGAGAAGAGAGCTTACAGCTTGT 1864
Db 1821 GGTATTAGCATGATGTGGGAATTAAGAAGATCATTTACAGAGGAGCGTTTAAAGTTGGTA 1880
QY 1865 TTGCTCAATCCTCTAAGTGAAGTGAAGAAACTGAAGAACTGAACAAATCGAAGT--TCCAAAT 1921
Db 1881 TTGTCAAATCCAAAGAGAGGTGCTGGAAGAAATTAACGATCCAAATTCATCGGTGAT 1940
QY 1922 CATTTAGGGAAGAAATGGATCTATCTGACTGTTTGAAGAGCCGTTGGAGCATGCAACTTC 1981
Db 1941 CATTTGGGCAAGAGTGGATGTTCTTAAACGTTAGGAGACGAGTGGAGGCTTGTAGCTAC 2000
QY 1982 AATCTAGTGCACGAACAAACGAACCC 2007
Db 2001 ATGCTTTCACACGTTTAAACCCGAACC 2026

RESULT 4
BNAS581745
LOCUS BNAS581745 2193 bp mRNA linear PLN 02-SEP-2003
DEFINITION Brassica napus mRNA for sulphate transporter (bat3.1 gene).
ACCESSION AJ581745
VERSION AJ581745.1 GI:34481597
KEYWORDS bat3.1 gene; sulphate transporter.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1
Buchner, P., Stuiver, E.E., Hawkesford, M.J. and de Kok, L.J.
Analysis of the effect of sulphur nutrition and atmospheric H₂S
exposure on the expression of sulphate transporter genes in Curley
cale
Unpublished
2 (bases 1 to 2193)
Buchner, P.
Direct Submission
Submitted (01-SEP-2003) Buchner P., Agricultural and Environmental
Division, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ,
UNITED KINGDOM

FEATURES
source Location/Qualifiers
1..2193
/organism="Brassica napus"
/mol_type="mRNA"
/variety="Starlight"
/db_xref="taxon:3708"
1..2193
/genes="bat3.1"
1..79
/genes="bat3.1"
80..2056
/genes="bat3.1"
/codon_start=1
/evidence=experimental
/product="sulphate transporter"
/protein_id="CAE4642.1"
/db_xref="GI:34481598"
/translation="MGTEENTPQGAEEPRRHHAVEAPDPPFLKSLQSVKETLFP
DPPFPKQNTTSRQVVLKYLPLEWAPLNFKLKPSDLIAGTISLASLTPQGIS
YAKLANLPILGLIYSFVPLVAVLGSSKDLAVGTVAVGSLLUTGAWLSKEVDREKDP

KLXLHLAFTATFPAGVLEASLGIFRLGFIWDFLSHATIVGFMGAAATVVSLOOLKGFIF
KLHFTAEADVISVMRSVFSQTHQWESVGLKGLNPLSVLDIVFTSPYNSTAKTKPLFWAA
MAPLTSVILGLLVYFTHAERGVQVIGDLKGLNPLSVLDIVFTSPYNSTAKTKGLI
TGIIILAEGLAVGESFAMEKNVINDGNKSMIAFGMNI VGSPTS CVLITTPSPRSVAVN
FNAGCKTAVSNVMAI VAMPTLLFHPFPHFPLVPLVLSLI VVAMGLGLIDYQAIHLW
KVDKDFVCMISAIFGVSVGSVEIGILVAVVISIARLLLVFVSPRPVAVGNIPNTMIY
RNTDQPYRISVIFGLJLILEIDAPIYFANAGYLREIRTRWVDEEDRIKASGNSLQYV
ILDSAVGNIDTSGISMBEIKKIMDRRELKVLVLANPKGEVVKLITRSKFIENLGKE
WMFLTVGEAVEACS FMLHTSKTEPASKERPNNV"
2057..-2180
/gene="bst3.1"

ORIGIN

Query Match 42.1%; Score 1032; DB 8; Length 2193;
Best Local Similarity 72.2%; Pred. No. 1.9e-245;
Matches 1357; Conservative 0; Mismatches 520; Indels 3; Gaps 1;

Qy	131	GTGAGGTTCCACGGCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACT	190
Db	143	GTGGAGGCTCCGGATCCTCAGCGGTTCTTGAAGTCACTTCAGTATTCAAGTGAAGAACT	202
Qy	191	TTCTTCCCTGATCACCCTTTGAGGAGTTCAAGAACCAAGCCAGCTTCCAGAGTTCATG	250
Db	203	CTGTCTCCAGACGATCCTTTCAGACAGTTCAAGAACCAAGACATCAAGACAAAGTTGTA	262
Qy	251	CTTGGGCTTCAGTTCCTTCTTCCCATTTTTCGAATGGGCTCCCAAAATACACCTTTTCAGTTC	310
Db	263	CTAGGCTCAAAATCTTCTTCCCAATATTGGAATGGGCTCCACTCTACAATTTCAAGCTC	322
Qy	311	TTGAAAGCTGACCTCATAGCTGGCATCAACATCGCTAGCTTTGGCCATTTCTCAGGGCATC	370
Db	323	TTTAAATCAGATCTCATCGCTGGAATCACCATCGCTAGCTCGCCATCCTCAGGGCATC	382
Qy	371	AGTTATGCCAAGCTCGCCACCTCCCTCCAATCTTGGACTATATTTCGAGCTTTATACCA	430
Db	383	AGTTACGCTAAACTCGCTAACTTGGCTCCCATCTTCTGGGCTTTTACTCGAGTTTGTTC	442
Qy	431	CCATTGATTTATCCGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGACTGTGGCGGTT	490
Db	443	CCATTGGTGTTCGCTGTGCTGGGAGTTCAAAGGAATTTGGCGTGGGACGCTGGCGGTT	502
Qy	491	GGATCGCTTCGATGGGTTCGATGTGAGTAATGCGCGTTGATCCCAATGAAGACCCAAAG	550
Db	503	GGATCTTTGTGACAGGTGCGATGCTGAGCAAGAAGTTGACGCTGAGAAGAATCTTAAG	562
Qy	551	CTTTACTCCACCTGCTTTCAGAGCTACATATTGCTGGTGTGTTTTTCAGGCTGCTTG	610
Db	563	CTCTACTCTCATCTGCTTTCATGCGCACGTTTTTTCGAGGGGCTTCTCGAAGCCTCTCTC	622
Qy	611	GGTCTGTTTAGGTGGGGTTGATCGTGGATTTTTCTGTCAATGCCAACCAATAATAGGGTTC	670
Db	623	GGCAATTTTCAGGTTTAGGTTTCATAGTGGATTTTTCTATCGCATGCAACGATAGTGGGATTC	682
Qy	671	ATGGAGGAGACGCCACGGTGTGTGTGACAGCAACTAAATCGAATCTTTGGCCCTTGAG	730
Db	683	ATGGAGGAGACGACGACGGTGTGTGTGACGCAACAGCTTGAAGGATTTTTTGGACTTAA	742
Qy	731	CATTTCCCATGAGCTGATATCATATCATATCAGTATGCGCTCTGTTTTTACCACCAACTCAT	790
Db	743	CATTTCCACAGAAGCCACTGACGTTATTTCTGTATGCGTTTTCAGTCTTCTCTCAAACTCAT	802
Qy	791	GAGTGGAGGTGGGAAAGTGTGTGTGTAGGATGTGTCTTCATTTTCTTCTCTCTTAGCACA	850
Db	803	CAGTGGAGATGGGAGATGGTGTCTCGGTGTGTGTCTTCTTCTTCTTCTTCTTCCACC	862
Qy	851	AGATACTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGACGAATGGGCCATTTGACG	910
Db	863	AGATATTTTCAGCAACGAAGAAACCAAAATTTCTTTTGGGTAGTGCATGGCTCCTTTGACC	922
Qy	911	TCCGTTTATATGGGAAGTCTCTTGGTGTATTTCACCTCAGCCGAGAACGACGCTGTGAA	970
Db	923	TCAGTGAITCTTGGAAAGCCTCTTGGTGTACTTTCATCTCAGCGGAGAGACATGGTGTTC	982

Qy	971	GTGATAGGAGAACTGAAGAAGGGTTTGAATCCACCATCATCTCACAATCTGGTATTGTG	1030
Db	983	GTGATAGGAGGACCTGAAGAAAGGGTTGAACCCGCTCTCTGTGTGATCTTGTCTTCACT	1042
Qy	1031	TGCGCTTACATGACTACAGCTGTCAAACTGGCATTTGTCGTTGGCATCATATCATCTGGG	1090
Db	1043	TGCGCTTACATGCTCAACAGCTCTCAAACTGGCCCTCATCACTGGGATCATATTCTTGA	1102
Qy	1091	GAAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATAAAAAATTACAATATTGATGSCAAC	1150
Db	1103	GAAGGATAGCTGTGGGAAGGAGCTTTGCCATGTTTCAAGAACTACAAATAGACGGTAAC	1162
Qy	1151	AAAGAGATCATAGCTATTGGGACCATGAACGAGTTGGTTTCTTTCACTCTTTGTACCTC	1210
Db	1163	AAAGAAATGATAGCGTTTGGAAATGATGAACATCGTTGGTTTCTTCACTCTTTGTACCTC	1222
Qy	1211	ACAACAGGACCAATTTTTCGCGTTTCGGCTGTGAACATTAACGCTGGATGCAAGACAGAGCT	1270
Db	1223	ACAACCTGGACCGTTTTTCAAGATCGGCGGTGAACATTTCAATCGCGGTTGCAAGACTGCCGTG	1282
Qy	1271	TCCAACATTTAATGTCACCTTCAGTAAATGTTGACATTTGTTATTTCTCTGACACCTTTGTC	1330
Db	1283	TCCAACATAGTAATGGCCATTCGGTTATGTTTCACTGCTCTTCCACAGGCTTTTTTTT	1342
Qy	1331	CATTACACTCCCTCGTGTGCTATCAGCTATTATCGTATCTGCAATGCTTGGACTCATA	1390
Db	1343	CACACAGCCCTCGTCTGCTCTCTCTCCATCATCATGTCGCAATGCTCGGACTCATC	1402
Qy	1391	GATTAAGAACGACCATCCATCTATTAAAGGTTGACAAATTTGACTTTTGGTGTGTCATG	1450
Db	1403	GACTATCAAGCTGCCATCCATCTTTTGAAGGTTGACAAATTCGATTTCTCTGCTGTCATG	1462
Qy	1451	AGTGATACATTTGGCGTGTCTTTGCGAGTGTGAAATTTGGCTTAGTCATAGCTATTGTA	1510
Db	1463	ACGCGCTACTTTGGGGTGTGTTGCTAGTGTGAGATCGGACTTTGTCGCAAGTGGTA	1522
Qy	1511	ATATCTGTAATTTTCGGGTACTCTTATTATTGAAGCCCAAGGACATTCGTTTTTGGGCAAC	1570
Db	1523	ATATCGATAGCAAGGTTGTGCTGTTGCTGTCGAGGCGGAGACTCGCGTGAAGGAAAC	1582
Qy	1571	ATTCCAATTTCTGTGATATACCGAAATGTTGAGCACTCAAAATGCAAAACATGTTCTT	1630
Db	1583	ATACCAAAACACCATGATCTATAGGAACACTGACCAAGTATCTTACTCAAGAAATTTGTTCT	1642
Qy	1631	GGAAATGCTAAATTTCTAGAGATTGATGCACCAATTTTACTTTGCCAATGCCAGCTATTAA	1690
Db	1643	GGTCTTCTCATCTCGAGATTGATGCTCCACTACTTTGCCACGCTGTTTACTTGGC	1702
Qy	1691	GAAAGGATCAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACAGGGGAGACT	1750
Db	1703	GAGAGAAATCACAAAGGTGGTTCGATGAAGAGGAAGATAGGATCAAGAGCATCAGGAGAAAT	1762
Qy	1751	AGTTTGCAGTATGTTTAAATGATATGATGCTGTTGGAAACATTTGATACATAGTGGAAATA	1810
Db	1763	AGTTTACAATATGTTTATATCGATATGTCGTGTTGTTGTAATATCAGACAGAGTGGTATA	1822
Qy	1811	AGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGACTACAGCTTCTTTGGTTC	1870
Db	1823	AGCATGATGGAGAAATTTAAGAAATTCATGGACAGAGAGAGTTAAGTTAGTATTGGCA	1882
Qy	1871	AATCTCTGAAGTGAAGTGAAGAACTGAACAAATCGAAGT---TCCAAATCATTTA	1927
Db	1883	AATCCAAAAGGAGAGGTCGTGAAGAAACCTAACAGATCCAAATTCATCGGTGAAAATTTG	1942
Qy	1928	GGGAAGAAATGATCTATCTGCTGTTGAAGGCGGTTGGAGCATGCACTTCAATCTA	1987
Db	1943	GGCAAGAGTGGATGTTCTTAACTGTTGGAAGCGCGTGAAGCTGTGATGTTTCATGCTT	2002
Qy	1988	CGTGAAGCAAAACGAAACC 2007	
Db	2003	CACACGTCAAGACCGAACC 2022	

RESULT 5	AX506807	AX506807	1977 bp	DNA	linear	PAT 27-SEP-2002
LOCUS	Sequence 1502 from Patent WO0216655.					
DEFINITION	AX506807					
ACCESSION	AX506807.1	GI:23388044				
VERSION						
KEYWORDS	Arabidopsis thaliana (thale cress)					
SOURCE	Arabidopsis thaliana					
ORGANISM	Arabidopsis thaliana					
REFERENCE	1	Harper, J.P., Kreps, J., Wang, X. and Zhu, T.				
AUTHORS						
TITLE	Stress-regulated genes of plants, transgenic plants containing same, and methods of use					
JOURNAL	Patent: WO 0216655-A 1502 28-FEB-2002;					
	The Scripps Research Institute (US) ; Syngenta Participations AG (CH)					
FEATURES	Location/Qualifiers					
source	1. 1977					
	/organism="Arabidopsis thaliana"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:3702"					
ORIGIN						
Query Match	41.6%;	Score 1018.2;	DB 6;	Length 1977;		
Best Local Similarity	72.3%;	Pred. No. 5.1e-242;				
Matches 1364;	Conservative	0;	Mismatches 518;	Indels 5;	Gaps 3;	
QY	125	CACCACTGAGGTTCCACCGCCACACAGCCGTTTTCAGTCTCTAAGTACTCTTTGAG	184			
DB	58	CACACGGTGGAGGCTCCACAACCTCAACCCGTTCTTGAAGTCACCTCAGTACTCAGTGAAG	117			
QY	185	GAGACTTCTTCCCTGATGACCTTTTGAGGAGTTCAAGAACCAAGCCAGCTTCCAAGAG	244			
DB	118	GAACTCTGTTTCAGAGACCTTTTGAACAATTTAAGAACCAAAATGATCAAGAAA	177			
QY	245	TTCAATGCTGGCCCTCAGTCTTCTTCCCAATTTTGGATGGGCTCCCAATACACCTTT	304			
DB	178	TTTGTGTTAGGCTCAAAATACTTCTCCCGATTTTCGAATGGGCACCAGCTACAATCTC	237			
QY	305	CAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATCTCTCAG	364			
DB	238	AAGTTCTTCAAAATCAGATCTCATGCGCGGAATCACCATCGCTAGCTCTGCGCATCCCTCAG	297			
QY	365	GGCATCAGTTATGCAAGCTCGCAACCTCCCTCCAAATCTTGCACTATATTCGAGCTTT	424			
DB	298	GGCATCAGTTATGCGCAAACTTGCTAATCTGCCCCCAATCTTGCCCTTATTCGAGTTT	357			
QY	425	ATACCAACCATTTATATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACTGTG	484			
DB	358	GTACCGCCATTGGTATACGCGGTGTAGGAGTTCAAGGACTTAGCGGTGGGAACGGTT	417			
QY	485	CGGTTGGATCGCTTCTGATGGGTTTCAATGTTGAGTAAATCGCGTTGATCCCAATGAAGAC	544			
DB	418	CGGTTGCGCTCTCTGTTGACAGGTGCGATCTGAGCAAGAAGTTGATGCTGAGAAGAT	477			
QY	545	CCAAAGCTTTTACCTCCACCTGGCTTTCACAGCTACATTTATTTGCTGTGTGTT-TTTCAGGC	603			
DB	478	CCTAAGCTTTTACCTTCACTTGGCTTTTCCGCCCAATCTTTCGCCGCGGTCTCGAAGCC	537			
QY	604	TGCCCTTGGGCTGTGTTTATGTTGGGTTGATCGTGGATTTTCTGTCAATGCAACCAAT	663			
DB	538	TCTCTTGGAAATTTTCAGGGTTAGGGTTTCATAGTGGAATTTCTATCGCATCAACGATAGT	597			
QY	664	AGGTTTCATGGGAGGACGCCAGCGGTGTGTCTGCAGCACTAAATTCGATTTCTTGG	723			
DB	598	AGGATTCATGGGAGGACGCCAGCGGTGGTGGTCTGCAACAGCTTAAAGGGTATTTTCGG	657			
QY	724	CCTTGACATTTTCAACCATGGAGCTGATATCATATCAGTCAGTCGCTCTGTTTTTCAACCA	783			
DB	658	ACTTAAACATTTTACAGACTCTACCGATGTTAICTCTGTATCGGTTCCGTTTTTCTCCCA	717			
QY	784	AACCTATGATGGAGGTGGAAAGTCTCTGTGTGTAGATGTCTTCATTTTCTTCTCTCT	843			
DB	718	AACCTACGAGTGGAGATGGAAAGTGGCGTCTCTTGGATGTGGTTTCTATTTCTTCTTCT	777			
QY	844	TAGCACAGATACTTCCAGCAAAAACGACCAAGGTTTTTTTGGGTGTGACGAATGCGGCC	903			
DB	778	CTCCACCAATATTTCAGCATCAAGAAACAAAATTTCTTTGGGTGGCGGATGGCTCC	837			
QY	904	ATTGACGTCGCTTATATTGGGAAGTCTCTTGGTTTATTTTCACTCACGCCGAGAGAACGG	963			
DB	838	TTTGACCTCAGTGATTTCTTGGGAAGTCTCTTGGTTTACTTCTCACTCACGCTGAGAGACATGG	897			
QY	964	TGTTGAAGTATAGGAACACTGAAGAGGTTTGAATCCACCATCACTCACAAATCTGGT	1023			
DB	898	TGTTCAAGTGATAGGGGACCTGAAGAAAGGGTTGAATCCACTCTCCGGTTCTGATCTCAT	957			
QY	1024	ATTTGTGTCGCTTACATGACTACAGCTGTCAAAACTGGCATTTGTGGTGGCATCATATC	1083			
DB	958	CTTTACTTCCCCTTACATGTCCACAGCTGTCAAACTGGCCTCATCTCTGGCATCATTC	1017			
QY	1084	ACTTGGGAAGGAATAGCAGTAGGAAGAGCTTTTGCATGTATATAAAATTTTACAATATTGA	1143			
DB	1018	TCTCGC-GAAGGAGTAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTACAACATAGA	1076			
QY	1144	TGGCAACAAAGAGATAGTACTATTGGGACCATGAACGTAGTTGGTTCTTTCACTCTTG	1203			
DB	1077	CGGGAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCATCGTTGGTTCTTCAATCTTG	1136			
QY	1204	CTACCTCAACAACAGGACCAATTTTCGGGTTTCGGCTGTGAACACTATAACGCTGGATGCAAGAC	1263			
DB	1137	TTACCTCAACAACGGAACCAATTTTCAAGTTCGGCAGTGAACACTCAACGCGGTTGCAAGAC	1196			
QY	1264	AGCAGCTTCCAACATTAATATGCTCACTTGCAGTAATGTTGACATTTGTTATTTCTCTGACCC	1323			
DB	1197	CGCAATGTCCAACATAGTATGCGGATTCGGGTTATGTTTACACTCTCTTCTCTCACACC	1256			
QY	1324	CTTGTTCATATACATCTCCCTGGTGTCTATCAGCTATTTATTCGATCTGCAATGCTTGG	1383			
DB	1257	GCCTTTTCACTACACCACTCGTCTCTCTCTGCCATCATATTCGCAATGCTCGG	1316			
QY	1384	ACTCATAGATTTAAGAGCAGCCATCCATCTATTTAAGGTTTGACAAATTTGACTTTGCTGGT	1443			
DB	1317	ACTCATGATCTCAAGCTGCCATCCATCTCTGGAAGTTTGAAGTTTCAAGTTTCACTTCTCGT	1376			
QY	1444	GTGCATGAGTGCATACATTTGGCGTGGTCTTTTGGCAGTGTGAAATTTGGCTTAGTCAATGC	1503			
DB	1377	CTGCATGAGCGCTACGTTTGGGTCGTATTTCCGCGTGTAGAGATTGGACTCGCTGTAGC	1436			
QY	1504	TATTGTAATATCTGTACTTCCGGTACTTCTATTTTATTTGCAAGCCCAAGGACATTCGTTTT	1563			
DB	1437	GGTGGCGATATCTATAGCGAGGTTGTGCTGTTGTGTCGAGGCCAAAAAATCGCGGTGA	1496			
QY	1564	GGCAACATTTTCAAAATCTGTGATATACCGAAATGTTGAGCACTATCAAAAATGCAAAACA	1623			
DB	1497	GGAAACATACCAACACGATGATCTATAGGAACACTGAGCAGTACCCATCATCAAGAAC	1556			
QY	1624	TGTTCTCGGAATGCTAAATTTCTAGAGATTGATGCACCAATTTTACTTTGCCAATCCGAGTA	1683			
DB	1557	CGTTCTCTGTAATCTCATCTTGAGATTGATGCTCCCATCTACTTTGCTAATSCCAGTTA	1616			
QY	1684	TTTAAGAAAGGATCACAAAGGTGGATTGATGAAGAAGAAGAAGATTTAAAGCTACAGG	1743			
DB	1617	CTTGGCTGAAAGAAATCATAAAGTGGATTGATGAAGAGAGAGAGAGTTTAAACAATCAGG	1676			
QY	1744	GGAGACTAGTTTGCAGTATGTTATTAATTCATATGAGTGTGTTTGGAAAAATTTCAATCAAG	1803			
DB	1677	AGAGACGCTTACAATATATTATCTCGATATGTCAGCTGTTTGGTAAATTCACACACAG	1736			
QY	1804	TGGAATAAGTATGCTTTGAAGAGGTGAAGAAGATTTACAGAGAGAGAGAGCTACAGCTTGT	1863			
DB	1737	CGGTATTAGCATGATGTTGMAAATTAAGAAAGTCAATTGACAGGAGAGCGTTTAAAGTTGGT	1796			

Qy	1864	TTTGGTCAATCTCTGAAGTGAAGTGAAGAACTGAAACAAATCGAAGT---TCCAAA	1920
Db	1797	ATTTGTCAAATCCAAAGGAGAGGTGCTGAAGAAATTAACCAAGATCCAAATTCATCGGTGA	1856
Qy	1921	TCATTTAGGGAAGAAATGGATCTATCTGACTGTCTCAAGAGGCCGTTGGAGCATGCAACTT	1980
Db	1857	TCATTTGGGCAAGAGTGGATGTTCTTAACGGTAGGAGACAGTCGAGGCTTGTAGCTA	1916
Qy	1981	CAATCTACGTGCAAGCAAAACGAACCC	2007
Db	1917	CATGCTTCAACAGTTTAAACCGAACCC	1943
RESULT 6			
AX651748	Sequence 589 from Patent WO03000898.		
LOCUS	AX651748	1977 bp	DNA linear PAT 22-MAR-2003
DEFINITION	Sequence 589 from Patent WO03000898.		
ACCESSION	AX651748		
VERSION	AX651748.1	GI:29154566	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE	1. 1977		
AUTHORS	Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M., Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.		
TITLE	Plant genes involved in defense against pathogens		
JOURNAL	Patent: WO 03000898-A 589 03-JAN-2003;		
FEATURES	Syngenta Participations AG (CH)		
	Location/Qualifiers		
source	1. 1977		
	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
ORIGIN			
Query Match	41.6%; Score 1018.2; DB 6; Length 1977;		
Best Local Similarity	72.3%; Pred. No. 5.1e-242;		
Matches 1364; Conservative	0; Mismatches 518; Indels 5; Gaps 3;		
Qy	125	CACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG	184
Db	58	CACACGGTGGAGGCTCCACAACTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAG	117
Qy	185	GAGACTTCTTCCCTGATGACCCCTTTGAGGCAGTTTCAGAAACAGCCAGCTTCCAAGAAG	244
Db	118	GAAACTCTGTTTCCAGACGACCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAA	177
Qy	245	TTCATGCTTGGGCTTCAGTCTTCTTCCCATTTTTCGAATGGGCTCCCAAAATACACCTTT	304
Db	178	TTTGTGTTAGGCTTCAAAATCTCTCTCCGATTTTCGAATGGGCACACGCTACAAATCTC	237
Qy	305	CAGTCTTGAAGAGCTGACCTCATAGCTGGCATCACCATCGTAGCTTGGCCATTTCTCTCAG	364
Db	238	AAGTCTTCAAAATCAGATCTCATCGCGGAATCAACCATCGTAGTCTCGCCATCCCTCAG	297
Qy	365	GGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTGGACTATATTCGAGCTTT	424
Db	298	GGCATCAGTTATGCCAAGCTTGTCTAACTTGGCCCCAAATTTCTTGGCCCTTTATTCGAGTTTT	357
Qy	425	ATACCAACCATTTGATGATGATGGT	

Qy	1142	GATGGCAACAAGAGATGATAGCTATTGGGACCATGAACGTAGTGTGGTTCTTTCACCTCT	1201
Db	1087	GATGGGAACAAGAGATGATAGCTTTGGGATGATGAACATTCTGGTTCTTCTCTCT	1146
Qy	1202	TGCTACTCTCAACAGACGACATTTTCGGTTCGGCTGTGACTATATACCGTGGATGAAG	1261
Db	1147	TGCTATCTCTACTACCGGGCCATTTTCACGTTTCGGCGGTGAACATACAAACGCGGGTTGCAA	1206
Qy	1262	ACAGCAGCTTCCAAACATTAATAATGTCACCTTGACAGTATGTTGACATTTGTTATTCCTGACA	1321
Db	1207	ACGGCGTTATCAACCGTGTGATGGCGTTGCAGTGGCGGTGACGTACTGTTCTTAAACG	1266
Qy	1322	CCCTTGTCATTAACACTCCCTCGGTGGTGTCTACAGCTATTATCGTATCGCAATGCTT	1381
Db	1267	CCGTTATTCTTCTACAGCCATTAGTAGTTCTGTCGTCGATCATATCGTCGTCATGCTA	1326
Qy	1382	GGACTCATAGATTATGAAGCAGCATCCATCTATTATTAAGTTGACAAAATTTGACTTTGGT	1441
Db	1327	GGCTTAGTCGACTACGAAGCAGCCATTCATCTTTGGAACTCGACAAGTTTCGATTTCTTTC	1386
Qy	1442	GTGTGCATGATGTCATACATTGCGCTGGTCTTTGGCAGTGTTCGAAATTTGCTTGTAGTCATA	1501
Db	1387	GTTTGTCTCTCTGCTTACTTTGGCGGTGTTTTTGGTACCATCGAAATTTGTCTCATCTCTC	1446
Qy	1502	GCTATTGTAAATCTGTACTTCGGGTACTCTTATTTATTGCAAGCGCAAGCAATTCGTT	1561
Db	1447	TCCGTGGGAATATCAGTGATGAGTTAGTGTGTTCTGTCGGAAGACCAAGATTATATGA	1506
Qy	1562	TTGGGCAACATTCGAAATTCGTGTATATACCGAAATTTGGAGCACTATCAAAATGCAAAA	1621
Db	1507	ATGGGAAACATTCGAAACTCAGAAATATATAGGAACATTTGAGCATTTACCCCTCAAGCCATC	1566
Qy	1622	CATGTTCTCTGGAATGCTAATCTCTAGAGATTGATGCACCAATTTACTTTTGCCTAATGCCAG	1681
Db	1567	ACTCGCTCTTCTCTCTCTATCTTCCATCTCCCATCGATGGTCTTATTTACTTCGCGAACTTACC	1626
Qy	1682	TATTTAAGAGAAAGGATCACAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACA	1741
Db	1627	TACTTCGGACAGAGAATTTGGAAGTGGATCGATGAAGAGAGAGATATAATTTGAGAACTAGC	1686
Qy	1742	GGGGAGACTAGTTTGCAGTATGTTTATAATTGATAGTGTCTGTTGGAAACATTTGATACA	1801
Db	1687	GGAGATATCAGTCTACAATATATTGACTCGACATGAGTGTCTGTGGGGAACATCGACAG	1746
Qy	1802	AGTGGAAATAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGCTACAGCTT	1861
Db	1747	AGCGGTATTAAGCATGCTTGAAGAACTCAACAAATCCTGGGAGTGAAGGCTTAAGCTA	1806
Qy	1862	GTTTTGGTCAATCTGTAAAGTGAAGTGAAGAAACTGAACAAATCGAAGTTTCCAAAT	1921
Db	1807	GTGATAGCGAAACCGGGAGCAGAGGTGATCAAGAAGCTAAGCAAGTCCACCTTCATTGAG	1866
Qy	1922	CATTTAGGAGAAATGGATCTATCTGACTGTTGAAGAGCGCTTGGAGCATGCAACTTC	1981
Db	1867	AGCATTTGGTAAGAAACGGAATTTATCTACCGGTGGCAGAGCGCTGCGAGTTTTCGATTTTC	1926
Qy	1982	AATCTACGTGCAAGCAAAACGAACCCAAAGAAAGATGAA	2020
Db	1927	ATGCTCCACACGGCTAAACCCGACTCGCGGTACCGAA	1965
RESULT 8			
AJ601439			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
AJ601439			
Brassica oleracea mRNA for sulfate transporter (LST3.2 gene).			
LST3.2 gene; sulfate transporter.			
Brassica oleracea var. acephala (kale)			
Brassica oleracea var. acephala			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.			

REFERENCE	1	Buchner, P., Stuijver, E. E., Hawkesford, M. J. and de Kok, L. J.
AUTHORS	2	Analysis of the effect of sulphur nutrition and atmospheric H2S
TITLE		exposure on the expression of sulphate transporter genes in Curley
JOURNAL	Unpublished	cale
REFERENCE	2	(bases 1 to 2219)
AUTHORS	Buchner, P.	
TITLE	Direct Submission	
JOURNAL	Submitted (24-OCT-2003) Buchner P., Agricultural and Environmental	
	Division, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ,	
	UNITED KINGDOM	
FEATURES	Location/Qualifiers	
source	1..2219	
	/organism="Brassica oleracea var. acephala"	
	/mol_type="mrna"	
	/variety="acephala"	
gene	/db_xref="taxon:3713"	
	1..2219	
	/gene="LST3.2"	
5' UTR	1..71	
	/gene="LST3.2"	
	/evidence=experimental	
CDS	72..2024	
	/gene="LST3.2"	
	/function="involved in plasma membrane sulfate transport"	
	/codon_start=1	
	/product="sulfate transporter"	
	/protein_id="CAE53112.1"	
	/db_xref="GI:37998858"	
	/translation="MKRALQYPQVETPPPPQPPKSLKNTLNEILFADDPFGKIRNES	
	KTSKIELGRHVPILWARGYNLDYKSDVIGTITIASLAIPGISVAQLANLPPI	
	LGLYSGFVPPMVYAIMGSSKOLACTVAVASILLTAMLGKEVSAVENPKLYLHLAFTS	
	TFPAGLMQTCGLLRLGLVELLSHAAIIGFWAGAAATVVCLOQLKGLGLSHFTSTSD	
	VSVLRSIFSQSPVWRWSGLCCFLFPITPKYISKRPKLFWSAWAPLVSVIFG	
	SLFVFLHAQHGIIIGELEKGINPPSIHLVFTSPYVMLAKTGIITGVLAALAGI	
	AVGRFAMKYNNIDGNKEMIAFGNMNIFGFSFSCYLTITGPFSSAVNNASCKTAVS	
	NVMAVAVATLLPLTLPFFVPLVLSIIIAAMLGLVDYEAAMNLRDKDFPVC	
	LSAPFVVGFTIEGLILSVGISVLRLLFLVGRPKIYVMGNIONTPVNIROYPOAT	
	TLSSIIILHIDGPIYFANSVYLRDRIGWIDEEEEKLKSSENSLOYIILDLASVNGNI	
	DTSGFLMLEEVENKILRRDLKLVIANPGAELMKLUSKSPFETIGDMWHLHTVAEAVS	
	ACDFVLHTAKPDSPEKISREPEFNV"	
5' UTR	2025..2219	
	/gene="LST3.2"	
	/evidence=experimental	
ORIGIN		
Query Match	35.3%;	Score 864.8; DB 8; Length 2219;
Best Local Similarity	66.4%;	Pred. No. 7.2e-204;
Matches 1257; Conservative	0;	Mismatches 632; Indels 3; Gaps 1;
Qy	125	CACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG 184
Db	96	CCCCAAGTAGAAACCCCTCCACCAACCTTCTTGGAAATCACTAAAGAACACTCTCAAT 155
Qy	185	GAGACTTCTTCCCTGATGACCCCTTTTGAGCGAGTTTCAAGAACAA---GCCAGCTTCCAG 241
Db	156	GAGATCCCTTTCCCGCAGATCCGTTTCGAAAAATCAGGAACGAATCGAAAACGTCGAA 215
Qy	242	AAGTTCATGCTTGGCCTTCAGTTCCTTCTCCCATTTTTCGAATGGGCTCCCAAAATACAC 301
Db	216	AAGATCGAGTTAGGGTTGAGACACGCTGTTCGCCGATCTTGGAAATGGCTCGTGGTTACA 275
Qy	302	TTTTCAGTCTTGAAGCTGACCTCATAGCTGCATCACCATCGCTAGCTTGGCCATTCCT 361
Db	276	TTGGACTACTTGAATCAGACGTTATCTCAGGGATCAGATAGCAAGTCTTCTATTCCT 335
Qy	362	CAGGCGATCAGTTATGCCAAAGCTCCGCCAACCTCCCTCCCAATTTCTTGGACTATATTCG 421
Db	336	CAAGGGATTAGCTATGCTCAACTCGCTAATCTCCCTCCAATCCTTGGTCTTTTACTCGAG 395
Qy	422	TTTATACCAACCATTTGATTTATCGGATGATGGGTAGCTCGAGGGATTGGCAGTGGGACT 481

```
Db 396 TTCTGTCGGCGGATGTTATACCGCATTAATGGGAAGTTCAAAAGATTTTAGCGGTGGGCACG 455
Qy 482 GTGCGGTTGGATCGCTTCTGATGGGTTGATGTTGATGATGCGGTTGATGCCAATGAA 541
Db 456 GTAGCGGTGGCGTGGCTGTTAAACGGCGGCAATGCTTGGAAAAGAGGTAAGCGCGGTGGAG 515
Qy 542 GACCCAAAGCTTACCTCCACCTGGCTTTCACAGCTACATTAATTTGCTGGTGTGTTTCAG 601
Db 516 AATCCAAACCTTACCTTACCTTAGCTTACCTTCCACTTCTTCTGCGCGGACTCATGCAA 575
Qy 602 GCTGCTTGGGTCTGTTTAGTGTGGGCTTGATCGTGGATTTCTGTACACATGCAACCAT 661
Db 576 ACCTGCCCTTGGCTCTTACGGTTAGGATTTCTGTGGAGATATTTGTCGACGCGGCGATC 635
Qy 662 ATAGGGTTATGGAGAGAGAGCCACCGGTGGTGTCTGAGCAACTAAAATCGATCTTT 721
Db 636 ATAGGGTTATGGGTGGAGAGAGCGACGGTGGTGTCTCAACAACACTGAAAGGATTAATT 695
Qy 722 GGCTTGGAGCATTTCCACCATGGAGCTGATATCATATCAGTGATGGCTCTGTTTTCACC 781
Db 696 GGCTCAGTCACCTTACCCCATTCACCTGATGTCGTCTCCGTTCTTCTGTTCCATCTTTCT 755
Qy 782 CAAACTCATGAGTGGAGGTGGGAAAGTGTCTGTAGGATGTCTTCAATTTCTTCTCCTC 841
Db 756 CAGTCTCTGTGGAGATGGAGAGTGGGTACTAGGGTGTCTTCTTCTTCTTCTTCTTCT 815
Qy 842 CTTAGCAAGAGATCTTTAGCAAAAAAGCAAGGTTTTTTTTGGGTGTGAGCAATGGCG 901
Db 816 CCTATACCAAGTACATTAGCAAGAGAGACCAAGCTATCTGGATATCTGCAATGGCT 875
Qy 902 CCATTGAGTCCGTTATATTGGGAAGTCTCTTGGTTTATTTTCACTCAGCGCGAGAGCAC 961
Db 876 CCATTGTTTCTGTCACTTTTGGAAAGCTCTTCTGTTTACTTCTCCATGCCCCAGTTCCAT 935
Qy 962 GGTGTTGAAGTGTAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAAATCTG 1021
Db 936 GGATCCAAATATTAGAGAAATTAGAGAAAGGATTAATCCACGTCATCACACATCTG 995
Qy 1022 GTATTTGTGTGCGCTTACATGACTACAGCTGTCAAACTGGCAATTTGCTGTGGCATCAT 1081
Db 996 GTTTTCACGTCACCTTATGTATGCTGCTGCTCTTAAACCGGCATTTATCACTGGAGTCTTA 1055
Qy 1082 TCATTTCGGAAGATAGCAGTAGGAGAAAGCTTTGCAATGTATAAATTTACAATTT 1141
Db 1056 GCTCTTGTCTGAAGAAATGTCAGTGGGAGAAAGCTTTGCAATGTATAAAGAAATTACAATA 1115
Qy 1142 GATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTCACTCT 1201
Db 1116 GATGGTAACAAAGAGATGATGCAATTTGGGATGATGAAATTTTGGTCTCTTCTCTCT 1175
Qy 1202 TGCTACTCAACAAGAGCAATTTTGGCGTTCGGCTGTGAACATATAACGCTGGATGCAAG 1261
Db 1176 TGCTATCTCACTACCGGACCGTTTTCACGTTTCGCGGTGAATACAAAGCCAGTTGCAAA 1235
Qy 1262 ACAGCAGCTTCCACATTAATGATGCTGATGATGATGATGATGATGATGATGATGATGAT 1321
Db 1236 ACAGCGGTATCAACCGGTGATGCGGTTGCGATGATGATGATGATGATGATGATGATGAT 1295
Qy 1322 CCCTTGTTCATTAACACTCCCTTGGTGTGCTATGATGATGATGATGATGATGATGATGATG 1381
Db 1296 CCATTAATCTTTTATACCGCGTTAGTAGTTCTGCTGATGATGATGATGATGATGATGAT 1355
Qy 1382 GGAATCATAGATTATGAAGAGCCATCCATCTATTTAAGGTTGACAAATTTGATTTG 1441
Db 1356 GGAATCTGCTACTATGAAGCAGCATGAATCTTTGGAGACTGACAAAGTTTCGATTTCTTC 1415
Qy 1442 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1501
Db 1416 GTATGCTCTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1475
Qy 1502 GCTATTGTAATATCTGATCTTCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1561
Db 1476 TCTGTGGGAATATCAGTCTCTGAGTTACTGTTGTTCTGTTGGGAAGACCAAGATTTATGTA 1535
```

```
Qy 1562 TTGGGCAACATTTCAAATTTCTGTATATACCGAAATTTTGAAGCACTATCAAAATGCAAAA 1621
Db 1536 ATGGGAAACATTTCAAGAACAGAGCCCTATAGGAACATTTAGAACAGTACCCTCAAGCCACT 1595
Qy 1622 CATGTTCTCGAATGCTAATTTCTAGAGATTTGATGACCAATTTTACTTTGCAATGCGAGC 1681
Db 1596 ACTCTCTCTAGTATCATCTTCTCCATATCGATGGTCTTATTTACTTTGCGCAACTCTAGC 1655
Qy 1682 TATTTAAGAGAAAGGATCAACAGGTGATGATGAAGAAGAAAGAAATTTAAAGCTTACA 1741
Db 1656 TACCTCGAGATAGATTTGGNAGTGTGATGAGAGAGAGAAAGAACTGAGAAAGAGT 1715
Qy 1742 GGGGAGACTAGTTTTCAGTATGTTAATTTGATATGAGTCTGTTTGGAAACATTTGATACA 1801
Db 1716 GAAGAAACAGTCTACAATACATCATCTTATTTGATGTTGTTGAGTCTGTGGGAAACATCGAC 1775
Qy 1802 AGTGGAAATAGTATGCTTGAAGAGTGAAGAGATTTACAGAGAGAGAGAGACTACAGCTT 1861
Db 1776 AGTGGTATAGAAATGCTCGAGGAAGTCAACAAATATCTGGGACGAAGAGACCTCAAGCTA 1835
Qy 1862 GTTTTGGTCAATCTCTGAAGTGAAGTGAAGAACTGAAGAACTGAAGTTCGAAAT 1921
Db 1836 GTGATAGCTAACCCAGGACAGAGTTGATGAAGAGCTAAGCAAGTCCAGTTCCTCCAGAA 1895
Qy 1922 CATTTAGGAGAAATATGATCTATCTGACTGTTTGAAGAGCCGTTGGAGCATGCAACTTC 1981
Db 1896 ACCATCGCAAGAGCTGGATTCATCTCAGGTAGCAGAAAGCCGCTCTCAGCTGTGATTC 1955
Qy 1982 AATCTACGTGCGACGAAAGAAAGCAACCCAAAGAA 2013
Db 1956 GTGCTCCACACAGCTAAACCGGACTCACCGGA 1987

RESULT 9
AX653455 LOCUS 1974 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 3325 from Patent WO03000898.
ACCESSION AX653455
VERSION AX653455.1 GI:29156269
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
REFERENCE 1
AUTHORS Chang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M.,
Katsiriri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 3325 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source 1..1974
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN
Query Match 34.5%; Score 846; DB 6; Length 1974;
Best Local Similarity 66.4%; Pred. No. 3.4e-199;
Matches 1231; Conservative 0; Mismatches 620; Indels 3; Gaps 1;

Qy 131 GTCGAGGTTCCACCGCCACAGCGCTTTTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACT 190
Db 55 GTGCCATGCGCGCGGAGCGCTTCTTGGAGACGCTGGGGGGAACATGAAGAGACA 114
Qy 191 TTCTTCCCTGATGACCTTTGAGGAGTTCAAGAAACAAGCCAGCTTCCAAAGATTC--- 247
Db 115 TTCTTCCGAGACGACCGGTTTCAAGGTTGCGCGGAGCGCGGTGCGCGCGCGCG 174
Qy 248 ATGCTTGGCTTTCAGTCTTCTTCTTCCCATTTTTCGAATGGGCTCCCAATACACCTTTCAG 307
```

Db 175 GCGCGCGCGCTCCGGTACGTGTTCCGTTTCATGGAGTGGCGCGCGTCTGATACACCCCTCGGC 234
Qy 308 TTCTTGAAAGCTGACCTCTAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAGGGC 367
Db 235 ACCCTCAAGTCCAGCTTCATCGCGGATCACCATTGCGACGCTCCGCCATCCCCAGGGC 294
Qy 368 ATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTTGGACATATATTCGAGCTTTATA 427
Db 295 ATCAGCTACGCCAAGCTCGCCAACTCCCTCCGCTCTCTATTCGAGCTTCGTG 354
Qy 428 CCACCAATATGATTTATCGGATGATGGTAGTCTCAGGGATTTGGCATGGGAGATGTGGCG 487
Db 355 CCGCCGCTGGTGTACGCGATGATGGGAGCTCAGGGACCTGGCGGTGGGACGGTGGCG 414
Qy 488 GTTGGATCGCTCTGATGGTGGTGGATGTTGAGTAATGCCGTTGATCCCAATGAAGACCCA 547
Db 415 GTGGCGTCTGCTGATCGGCTCGATCGGTCGATCGAGCGAGAGGTGTCGGCGGGAGGACCCG 474
Qy 548 AAGCTTTACCTCCACCTCGCTTTTCACAGCTTACATTTATTTGCTGGTGTCTTTTCAGGCTGCC 607
Db 475 GCGCTGTACTGACGCTCGCGCTCACCGCACCTTCTTCGCGCGGTGTTCCAGGGCGCTG 534
Qy 608 TTGGGTCTGTAGGTGGGTGGATGCTGGATTTTCTGTTCATGCAATGCAACCAATATAGGG 667
Db 535 CTGGGGTCTTGAGGTGGGGTTCATCGTGGACTTCTGTGCGACGCCACCATCGTCGGG 594
Qy 668 TTCATGGGAGGACGACGCGGTGGTGTCTGCAGCAACTAAATCGATTCTTTGGCCCTT 727
Db 595 TTCATGGGCGGCCGCCACCGGTGGTGTCTGCAGCAGCTCAAGGGCATGTTTCGGCCCTC 654
Qy 728 GAGCATTTTCAACCATGGAGCTGATATCATATCAGTGTATGCGTCTCTTTTTCACCCAAACT 787
Db 655 GACCACTTTCACCAACCCGACCGACCTCGTCTCGTCTATGAGCTCGCTTCTTCCCAACCC 714
Qy 788 CATGAGTGGAGTGGAAAGTCTGTGTAGGATGTCTTCAATTTCTTCTCTCTTAGC 847
Db 715 CACCTCTGGCGATGGGAGAGCGTCTCATGCGGTGCGGCTTCTCTTCTCTCTCATC 774
Qy 848 ACAAGATACCTCAGCAAAAACGACCAAGTGTCTTTTGGGTGTCAGCAATGGCGCCATTG 907
Db 775 ACCCGCTTCTCAGCAAGAGGAGGCCAAGTCTTCTGGGTATCTGCAGCTGGCCATTG 834
Qy 908 AGCTCGCTTATATTTGGGAAGTCTCTTTGGTTTATTTTCACTCAGCCGAGAACGCGTGT 967
Db 835 GCGTCTGTCATCATCGGGAGCTCTCTGTGTACTCTCACATGCTGAAACCATGSCATT 894
Qy 968 GAAGTATAGGAGAACTGAAAGAGGTGTAATCCACCATCATCAAAATCTGTATTT 1027
Db 895 CAAGTATTTGGTTACCTGAAGAGGGCTGAAACCCCGTCTGCCACAAGCCCTGAACTTC 954
Qy 1028 GTGTGCGCTTACATGACTACAGCTGCAAAACTGGCATTGTCTTGGCATCATATCATT 1087
Db 955 TGTGCGCGTACATGATGTGCGCCCTGAAGACGGGATCATCACTGGCGTCAATGCGCCCTC 1014
Qy 1088 GCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAATATACAAATTTGATGCG 1147
Db 1015 GCTGAGGGGATGCGCTGAGGAGGAGCTTTGCAATGTCTCAAGAAATTTACCACATCGACGA 1074
Qy 1148 AACAAAGAGATGATAGCTATTTGGGACCATGAAGTAGTTGGTTCTTTACCTCTTGCTAC 1207
Db 1075 AACAAAGGATGATGCGGTTTGGACGATGAACATTTGGGGATCACTACCTCTTGTCTAC 1134
Qy 1208 CTCACAAAGGAGCAATTTTTCGCGTTCGGGTGTGAATGATATAAAGTGTGAAGACGAGCA 1267
Db 1135 CTCACACCGGTCTCTCGAGTTCGGCGGTCACTACAAACCGCGGTGCAAGACGGCG 1194
Qy 1268 GCTTCCAAATATATGTCATTTGAGTAATGTTGACATGTTGATTTCTGACACCCCTTG 1327
Db 1195 ATGTCCAAAGTGTATGTCGTTGGTGGGTGATGATCACGCTGCTGTCTTCTGACGCGCTG 1254
Qy 1328 TTCCATTACATCCCTGCTGGTGTCTATGACTTATTCATGCTATCTGCAATGCTTGACTC 1387
Db 1255 TTCCACTACACGCCGCTGGTGGTGTCTGCGGATCATCATGTGCGCGGATGCTGGGGCTG 1314

Qy 1388 ATAGATTATGAAGCAGCCATCCATCTATTTTAAGGTTGACAAATTTGACTTTTGTGGTGTGC 1447
Db 1315 ATCGACTACCCGCGCGCTCCACTGTGGAGGTGGACAGGTGACTTTCGCGTCTGC 1374
Qy 1448 ATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTACTCATAGTATT 1507
Db 1375 CTCGGCGCTACTCTCGGCGTCTCTTCGCGACGCTCGAGATCGGCTCGTTCGCGCTC 1434
Qy 1508 GTAATATCTGTAATCTCGGTACTTCTATTATTTGCAAGGCCAAGNACATTCGTTTGGGC 1567
Db 1435 GGGATCTCCATCTCTCGGCTGTCTGTTTGGCGCGCGCGAGGACGACGCTGCTCGGG 1494
Qy 1568 AACATTTCCAAATTTCTGTGATATACCGAATGTTTGAGCACTATCAAAATCCAAAACATGTT 1627
Db 1495 AACATCCCCACACATGATCTACCGGGATGGACAGTACACCGCGCGCAGAGGGTTC 1554
Qy 1628 CTGGAATGCTAAATCTAGAGATTGATGCACCAATTTTACTTTGCCAATGCCAGCTATTTTA 1687
Db 1555 CCGCGCTGCTGTCTCGCGTCTGACTCGGCCATCTACTTCCACCAACGCCAGTACCTG 1614
Qy 1688 AGAGAAAGATCACAAAGTGGATTGATGAAGAAGAAGAATTTAAAGCTACAGGGAG 1747
Db 1615 CGTGAGAGATCGCGCGTGGATCGACGACGAGGAGGACCAAGTGAAGGAGAGGCGCAG 1674
Qy 1748 ACTAGTTTGCAGTGTATTATTAATTTGATATGATGCTGTGTTGGAAACATTCATCAAGTGA 1807
Db 1675 ATGGGCATTCAGTACGTCTCTCGACATGGGTGCGTTCGGAAGCATCGACACCGCGG 1734
Qy 1808 ATAAATGCTTTGAAGAGGTGAAGAGATTACAGAGAGAGAGACTACAGCTTGTGTTTG 1867
Db 1735 ACGAGCATGCTGGATGAACTCAGGAAGACCTCGACAGGAGGGCTTCAGATTGTGTTG 1794
Qy 1868 GTCATCTCTGAAGTGAAGTGAAGAACTGAACAAATCGAACATTCCAAATCATTTA 1927
Db 1795 GCGAACCCTGGGAGCGAGATCATGAAGAAGTTGGACAGCTCCAAGGTGTTGAGGCGATC 1854
Qy 1928 GGGGAAGAATGATCTATCTGACTCTTGAAGAGGCGGTGGAGCATGCAACTTC 1981
Db 1855 GGCCATGATGGATCTTCCCAAGTGGCGAGGCGGTGGCGAGTGGCATTC 1908

RESULT 10
AK066932
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AK066932 2327 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone.J013093F02, full insert sequence.
AK066932
FLI CDNA; CAP trapper.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
1
The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team.; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group.; Oono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niihara,J., Ikeda,T., Ishibiki,J., Kawamura,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN., Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Yoshino,M. and Hayaishizaki,Y.

TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL	Science 301 (5631), 376-379 (2003)
MEDLINE	22752273
PUBMED	12869764
REFERENCE	2 (bases 1 to 2327)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayaishi, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, P., Hotta, I., Iida, J., Ikeda, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanaoka, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numaaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Ootani, N., Ota, Y., Ootani, Y., Ryo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shihashi, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
TITLE	Direct Submission
JOURNAL	Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice. URL : http://cdna01.dna.affrc.go.jp/cDNA/ NTAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M. FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryo, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, P., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numaaki, R., Ohno, M., Ootani, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayaishi, Y. Location/Qualifiers 1. 2327 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /clone="J013093F02"
FEATURES	Source
ORIGIN	Query Match 34.5%; Score 846; DB 8; Length 2327; Best Local Similarity 66.4%; Pred. No. 3.4e-199; Matches 1231; Conservative 0; Mismatches 620; Indels 3; Gaps 1;

```
QY 1208 CTCACACAGGACCATTTCGCGTTCGGCTGTGAACATATAACGCTGGATGCAAGACAGCA 1267
Db CTACACACGGTTCGCTTCAGAGTCGGCGTCAACTACAAAGCCGGCTGCAAGACGGG 1305
QY 1268 GCTTCCAAATTAATATGTCATTCGAGTATGTCAGATGTTGACATGTTATTCCTGACACCCCTTG 1327
Db ATGTCGAACGTGATCATGTCGGTGGCGGTGATGATCACGCTGCTGTGTTTGACCGCGCTG 1365
QY 1328 TTCATTTACACTCCCTCGTGGTGGCTATCAGCTATTATTCTGATCTGCAATGCTTGGACTC 1387
Db TTCACACTACACGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1425
QY 1388 ATAGATTATGAAGACAGCCATCCATCTATTATTAAGGTTGACAAATTTGACATTTGFGTGTGC 1447
Db ATCGACTACCCCGCGCGCTCCACCTGTGGCAGGTGACAAAGGTGGACTTCTCGCTCTGC 1485
QY 1448 ATCAGTGCATACATTGGCGTGGTCTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTATT 1507
Db CTCGGCGCCTTACCTCGCGGTCGTCTTGGCAGCGCTCGAGATCGGCGCTCGTCTCGCGCTC 1545
QY 1508 GTAATATCTGACTTCGGGTACTTCTATTATTATTCGAAGGCCAAGGACATTTGTTTTGGGC 1567
Db GGGATCTCCATCTCTCGCGTCTGCTGTTGCTGGCGGCGGAGGACGAGCTGCTCGGG 1605
QY 1568 AACATTCCAAATTTCTGTGATATACCGAAATGTTGAGCATTATCAAAATGCAAAACATGTT 1627
Db AACATCCCAACACGATGATCTACCGCGGATGGACCCAGTACACCGCGCGCAGAGGGTC 1665
QY 1628 CTTGGAAATGCTAATTTCTAGAGATTGATGACCAATTTACTTTTGGCAATGCCAGCTATTTA 1687
Db CCGGCGGTGCTGCTGCTCGCGTCTGCTGTTGCTGGCGGCGGAGGACGAGCTGCTCGGG 1725
QY 1688 AGAAGAAGATCAACAAGTGGATTGATGAAGAAGAAAGAAATTAAGCTACAGGGGAG 1747
Db CGTGAGAGGATCGCGCGTGGATCGACGACGAGGAGCAACAGTGCAGAGAGGGCGAG 1785
QY 1748 ACTAGTTTGCAGTATGTTAATTTGATATGAGTGTCTTTGGAAACATTTGATACAACTGGA 1807
Db ATGGGCAATTCAGTACGTCTGCTCGACATGGGTGCGGTGCGAAGCATCGACACCGCGG 1845
QY 1808 ATAGTATGCTTGAAGAGTGAAGAGATTACAGAGAGAGAGAGTACAGCTTGTGTTTG 1867
Db ACAGCATGCTGGATGAATTCAGGAAGACCCCTGGACAGAGGGGCGCTTCAGATTTGTGTTG 1905
QY 1868 GTCAATCTCTTAAGTGAAGTGAAGAAACTGAACAAATCGAAGTTCCAAATCATTTA 1927
Db GCGAACCCGGGAGCGAGATCATGAAGAAGTTGACAGCTCCAGGTGCTTGAGGCGATC 1965
QY 1928 GGAAGAAATGGATCTATCTGACTGTTGAAGAGCGGCTTGGAGCATGCAACTTC 1981
Db GGCCATGAGTGGATCTTCCCAACGGTGGCGAGCGGTGGCGAGTGGCACTTC 2019

RESULT 11
LOCUS AR477457
DEFINITION Sequence 3 from patent US 6696292.
ACCESSION AR477457
VERSION AR477457.1 GI:47234915
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1981)
AUTHORS Allen, S.M., Falco, S.C. and Thorpe, C.J.
TITLE Genes encoding sulfate assimilation proteins
JOURNAL Patent: US 6696292-A 3 24-FEB-2004;
FEATURES
    Location/Qualifiers
        source
            1..1981
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN
```

```
Query Match 31.2%; Score 765.2; DB 6; Length 1981;
Best Local Similarity 65.8%; Pred. No. 4.3e-179;
Matches 1112; Conservative 0; Mismatches 578; Indels 0; Gaps 0;

QY 313 GAAAGCTGACCTCAGCTAGCTGGCATCACCATCGCTAGCTTGGCCATTCTCAGGCGATCAG 372
Db 7 GGAGTCGCGACCTGATCGCGCGCATCACCATCGCCAGCCTCGCATCCCGAGGCGATCAG 66
QY 373 TTATGCCAAGCTCGGCAACCTCCCTCCATTCTTGGACTATATTGGAGCTTTTATACCAAC 432
Db 67 CTACGCCAAGCTCGCCACACTCGCCCGCTGCTCGACTCTACTCGAGCTTCTGTCGCGGC 126
QY 433 ATTGATTTATGCGATCATGGGTAGCTCGAGGGAATTTGGCAGTGGGACTGTGCGCGTGG 492
Db 127 GCTGTGTACGCGCTGATGGGGAGCTTCAAGACCTGGCGGTGGGACGCTGGCGGTGGC 186
QY 493 ATCGCTTCTGATGGGTTCGATGTTGAGTATGCGGTTGATCCCAATGAAGACCCAAAGCT 552
Db 187 GTCGCTGCTCATCAGCTCCATGCTCGGCAGCAGGTTGTCGCGACGAGAACCCCGTGTCT 246
QY 553 TTACCTCCACCTGGCTTTTCCACAGCTACATTTATTGCTGGTGTGTTTTTCAGGCTGCTTGG 612
Db 247 CTACCTGCACTCGCTTCCACGCCACTTCTTTCGCGGCGCTTCTCCAGGCTTCTGCTCG 306
QY 613 TCTGTTTAGTGGGTGATCGTGAATTTTCTGTCAATGCAACCATATAAGGTTTCAAT 672
Db 307 CCTCTCAGTGGGCTTTCATCGTGGACCTGCTGCGACGCGACGATCGTGGGGTTTCAAT 366
QY 673 GGGAGGACGACCGAGTGGTGTGTCGAGCAACTAAATCGATTCCTTGGGCTTGGAGCA 732
Db 367 GGGCGGCGGCGAGCGGTGTGCTGCGACAGCTGAAGGGCATGCTGGGCTTCTGCTCCA 426
QY 733 TTTTACCCATGAGCTGATATCATATCAGTGATCGCTCTGTGTTTTTCAACCAAACTCATCA 792
Db 427 CTTTACACCTCCACCGAGCTGCTTCCGTATGAATCCGTTTCAGCAGACACACCA 486
QY 793 GTGGAGTGGGAAAGTGTGTTAGGATGTGTTCTTCAATTTTCTTCTCTTTAGCACAAG 852
Db 487 GTGGCGTGGGAGAGCGTCTCTCGGCTGGGCTTCTCTTCTTCTCTCTCTCTCTCACCG 546
QY 853 ATACTTTCAGCAAAAACGACCAAGTTTTCCTTGGGTGTGAGCAATGGGCGCATTTGACGTC 912
Db 547 CTTTATCAGCAAGAGCGCTCCCAAGCTGTTCTGAGATCTCCGCGCGCGCGCTTGAACGTC 606
QY 913 CGTTATATTGGGAAGTCTCTTGGTTTATTCTCACTCAGCGCGAGAGACGCGTTTGAAGT 972
Db 607 CGTGTGTCGGAGAGGTTCTGTTGTACTCTCAGCAGCGCTGAAACACCGCATCGAAGT 666
QY 973 GATAGAGAACTGAAGAAGGTTTGAATCCACCATCACTCAAAATCTGGTATTTGTGTGTC 1032
Db 667 GATCGGTTACCTGAAGAAGGCTGAATCCACCGTGGTGAACGCTGCAATTTCTCAC 726
QY 1033 GCCTTACATGACTACAGCTGTCAAAACTGGCATTTGCTGGCATCATATCTTGGCA 1092
Db 727 GCGCTTACATGATGCTCGGCTCAAGACTGGGATCATACCGGCGCTCATTTGGCCTCGCGA 786
QY 1093 AGGAATAGCAGTAGGAAGAAGCTTTTGCATATGATATAAAATTTACAATATTGATGGCAACA 1152
Db 787 AGGAATCGCGCTGGGAGGAGCTTCGCCATGTTTCAAGAACTACCATGACGACGACCAACA 846
QY 1153 AGAGATGATAGTATTGGGACCATGAACGTAGTGGTTCCTTTTCACTCTTGTGTAACCTCAC 1212
Db 847 GGAGATGATCGCGATCGGAGCGATGAAGCTCTCTGGGCTCGCTCAGCTGCTGTGTAACCTGAC 906
QY 1213 AACAGGACCAATTTTCGCGTTCGGCTGTAACATATAACGCTGGATGAGACAGACGAGCTTC 1272
Db 907 CACGGGGCGCTTCTCGCTCCGCGTGAACCTTACAAACGCGGCTGAGACGCGCATGTC 966
QY 1273 CAACATTAATGTCACCTTGCAAGTAATGTTGACATTTGTTATTTCTTGACACCCCTTGTCCA 1332
Db 967 GAACGTGTCATGCTCGCTGGCGTGAATGCTCAAGCTGCTGTTCTTCTGACGCGCTGTCCA 1026
QY 1333 TTACACTCCCTGGTGGGTGATATCATGTTATTCGTAATCTGCAATGCTTGGAGATCATAGA 1392
```

Db 1027 CTACAGCGCTGGTGGTCTGTCGGCGATCATCGTCTCCGCGATGCTGGGCCCTGGTCGA 1086
Qy 1393 TTATGAAGCAGCCATCCATCTATTAAAGTTTGACAAATTTGACTTTTGGTGTGCATCAG 1452
Db 1087 CTTGGGCGCGCGCTGCACCTGTGGCGGTTCGACAAAGGTTCGACTTCTGCGTCTGGCGCG 1146
Qy 1453 TGCATACATTCGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTATTGTAAT 1512
Db 1147 CGCGTACCTGGCGTGGTCTTTGGCAGCGTTCGAGGTTCGGCGTGGTGGTGGTGGTGGTGG 1206
Qy 1513 ATCTGTACTTCGGGTACTTCTATTATTATGCAAGCCCAAGACATTCGTTTTGGGCAACAT 1572
Db 1207 CTCCTGTCTCGCGTCTGCTGTTCGTCGCCGCGCCAGGACACAGGTGCTCGGCAACAT 1266
Qy 1573 TCCAAATTCGTGTATATACCGAAATGTTGACACTATCAAAATGCAAAACATGTTCTCTGG 1632
Db 1267 CCCGGCACCATGTTGACCGGAGATGGACCACTAGTACGCCGCGGAGCGGTGGCGCG 1326
Qy 1633 AATGCTAATTCAGAGATTGATGACCAATTTACTTTGCCAATGCCAGCTATTAAAGAGA 1692
Db 1327 CGTCTCGTGTGGCGTGCAGCGCGCGCTACTTCCGCAACGCCAGCTACCTGCGAGA 1386
Qy 1693 AAGGATCAAAAGGTGGATTGATGAAGAGAGAGAGAAATTAAGCTACAGGGAGACTAG 1752
Db 1387 GAGGATCTCGCGGTGGATTCGACGACGAGAGAGAGCGCACCAAGAGCCAGGGCGAGATGG 1446
Qy 1753 TTTGAGTATGTTATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1812
Db 1447 CGTGGGTACGTTCTCGACATGGGTGCCATCGGTAGCATCGACAGCGGAGCGAG 1506
Qy 1813 TATGCTTGAAGAGGTGAAGAGATTACAGAGAGAGAGAGCTACAGCTTGTGTTGGTCAA 1872
Db 1507 CATCTGAGCAGCTCAACAGTCTCTTGGACAGAGAGGGAATGAGATCGTGTGGCGAA 1566
Qy 1873 TCCTGTAAGTGAATGATGAGAACTCAACAAATCGAAGTTTCAAAATCATTTAGGGAA 1932
Db 1567 CCCGGGAGCGAGATCATGAAGAAGCTGGACAGCTCCAAGGTGCTGGAGCAGATCGGCCA 1626
Qy 1933 GAAATGGATCTATCTGACTGTTGAAGAGCGGTTGGAGCATGCAACTTTCAATCTACGTGC 1992
Db 1627 CGAGTGGGTGTTCCGAGCGTGGCGAGCGGTGGCGTGGTGGTGGTGGTGGTGGTGGTGG 1686
Qy 1993 AAGCAAAACG 2002
Db 1687 GCACAAGCGC 1696

RESULT 12
AR477463
LOCUS AR477463 2067 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 15 from patent US 6696292.
ACCESSION AR477463
VERSION AR477463.1 GI:47234921
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2067)
AUTHORS Allen,S.M., Falco,S.C. and Thorpe,C.J.
TITLE Genes encoding sulfate assimilation proteins
JOURNAL Patent: US 6696292-A 15 24-FEB-2004;
FEATURES Location/Qualifiers
source 1..2067
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 29.4%; Score 720.6; DB 6; Length 2067;
Best Local Similarity 62.2%; Pred. No. 5.3e-168;
Matches 1134; Conservative 0; Mismatches 689; Indels 0; Gaps 0;
Qy 143 CCGCCACAGCGGTTTTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACTTTCTTCCCTCGAT 202

Db 8 CCACACAGACCACTCCCAAACTTAGGCACAGAGTCTCCGAAATCTTCTTCCAGAT 67
Qy 203 GACCCCTTTGAGCGAGTTCAAGAACCAAGCCAGCTTCCAAAGTTCATGCTTGGCCTTCAG 262
Db 68 GACCCCTCCACCGTTTCAAGAACCAAACTCGCTTTAAAGTTCCTCTCGACTTCAG 127
Qy 263 TCTCTTCTCCCACTTTTCGAATGGGCTCCAAATACACCTTTTCAGTGTCTTGAAGCTGAC 322
Db 128 TATCTCTTCCCACTTTTCGACTGGGCCCAAACTACAATCTTACCCCTCTCCGCTCTGAC 187
Qy 323 CTATAGCTGGCATCAACATCGCTAGCTGGCCATTCCTCAGGCATCAGTTATGCCAAG 382
Db 188 CTATCTCTGGCCTCAACATTCGCAGCTCGCCATTCCTCAGGAAATCAGTTATGCCAAG 247
Qy 383 CTGGCCAACTCCCTCCAAATTTTGGACTATATTTCCAGCTTTTATPACCACCATTCATTTAT 442
Db 248 CTTGGCCAACTTGCACCTATTTCTTGGATATATTTGGATTTTGTTCCTCCCATTCATATAC 307
Qy 443 GCATAGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTGTGGCGTGTGATCGCTTCTG 502
Db 308 TCGCTGTCTGAAAGTTCTAGACATCTGTGTGTGGACCTGTTCCATTTGGTCTTTGGTC 367
Qy 503 ATGGGTTGATGTTGAGTAAATGCGGTGATCCCAATGAAGACCCAAAGCTTTTACCTCCAC 562
Db 368 ATGGGATCAATGTTAAAGTGATNAAAATTTCTTACACTCAAGAACCTATTTCTTATCTGGGA 427
Qy 563 CTGGCTTTACAGCTACATTTATTTGCTGGTGTGTTTTTTCAGGCTGCTTGGGTCTGTTTGG 622
Db 428 TTGGCTTTACCCGCACTTCTTTTGGTGTGTTTTCAGAGTATTCAGAGTCTCTGGGTATATTAAG 487
Qy 623 TTGGGTTGATCGTGGATTTTCTGCATGCAACCATATATAGGTTTCATGGGAGGAGCA 682
Db 488 CTAGGCTTCTGTAATTTGATTTTCTGTCGAGGCAACGCTGTTGGATTTTCACAGGCGGTGCT 547
Qy 683 GCCACGGTGTGTCTGACGCAACTAAATTCGATTTCTGGCCCTTGAGCAATTCACCCAT 742
Db 548 GCCATATTATGTTGTCACCTGCGAGCAGCTGAAAGGTTTACTTTGGAATAGTGCACTTTTACCAGC 607
Qy 743 GGAGCTGATATCATATCATGATGCTGCTGTTTTTTCACCCAACTCATAGTGGAGGTGG 802
Db 608 AAGATGCAATTAATTCAGTAAACATCTCTGTTTTTCAAGCAAGACACAGAGTGTCTATGG 667
Qy 803 GAAAGTGTGTGTAGGATGTGTTTCATTTTCTCTCTAGCACAAGATCTTCAGC 862
Db 668 CAAACCAATCTTTTGGGATTCGGCTTCTTGGTCTTCTGCTGACAAAGGCACTTAGT 727
Qy 863 AAAAAAGCAGCAAGTTTGTGGGTGTGAGCAATGGCGCATTCAGCTCCGTTATATTG 922
Db 728 TTGAGGAAACCAAACTATTTCTGGGTTTCAGCAGCTGCCCCCATTTGACATCAGTTATTTCTG 787
Qy 923 GGAAGTCTCTTGGTGTATTTTCACTCAGCGGAGAACGCGTGTGGAAGTATAGGAGAA 982
Db 788 TCAACCAATTTAGTCTTTCTCTGGAATAAAGACTCATCAAAATTCAGTTATTTGGGCAC 847
Qy 983 CTGAAGAAGGTTTGAATCCACCATCACTCAAAATCTGGTATTTGTGTGCCCTTACATG 1042
Db 848 TTACCAAAAGGAGTTAATCCACCATCAGCAAAACATGTTATCTTCAATGCTCTTACTTG 907
Qy 1043 ACTAGAGTGTCAAACTGGCATTTGCTGGCATCATATCATCTTGGGAGGAAGAAATAGCA 1102
Db 908 GGTCTTGTCTATCAAACTGGCATCATCAGAGGATCTTATCTCTACTGAAAGAAATGCA 967
Qy 1103 GTAGGAAGAAGCTTTGCAATGTATAAAATTAACAATTTGATGGCACAAAGAGATGATA 1162
Db 968 GTAGGAGAACATTTGCTCTTCACTTAAGAACTACAGGTGATGGAACAAAGAAATGATG 1027
Qy 1163 GCTATTGGGACCATGAACGTAGTTGTTCTTTTCACTCTTGTCTTACCTCAACAGGACCA 1222
Db 1028 GCCATTTGGTCTAATGAACATAGCTGGCTGTTCTTCTCATGTTATGTTTACACGGGATCC 1087
Qy 1223 TTTTCGGTGTGGCTGTGAACTATAACGCTGGATGCAAGACAGAGCTTCCCAACATTTATA 1282

Db	1088	TTTTCTCGATGGCGTGTAACTATAATAGTCTGGAGCACAGACAACAGTTCCTCAATATAATC	1147
Qy	1283	ATGTCACCTTGACGTAATGTTGACATGTTATATCTGACACCTTGTTCCATTAACATCC	1342
Db	1148	ATGGCTGACGCTGTTAGTGACACTTCTGTTTCTCATGCTCTTCTTCTACTATACACCA	1207
Qy	1343	CTGGTGTGTATACAGCTATTATCTGTAATCTGCAATGCTTGGACTCAATAGATTAAGACA	1402
Db	1208	AATGTTGCTTAGCGGCATATATCATCTGCTGTGATTTGGTCTTAATAGATTAATCAATCT	1267
Qy	1403	GCCATCCATCTATTAGGTTGACAAATTTGACATTTGTGTGTGTCATGAGTGACATCAAT	1462
Db	1268	GCATATAAATTTGGGAAGTTGACAACTTGATTTCTTGGCCCTGTTGTGCTCCTTTTTT	1327
Qy	1463	GGCGTGTCTTTGGCAGCTGTTGAAATTTGGCTTAGTCTATAGCTATTGTAATATCTGACTT	1522
Db	1328	GGGGTCTGTTTCAATTTTCAGTGCCTTTAGTCTTGGTATAGCGGTTATCATATCAGTCC	1387
Qy	1523	CGGCTACTTCTATTATTGCAAGGCCAAGGACATTCGTTTGGGCAACATTCCTCAAAATTC	1582
Db	1388	AAGATCCTGCTTCTATGTCACCTCGACCAACACACTTTGGTTTGGGGAATATACCGGAACA	1447
Qy	1583	GTGATATACCAATGTTGAGCACTATCAAAATGCAAAACATGTTCTCGAATGCTAAAT	1642
Db	1448	CAAAATTTCCACACATATAACCAATACAAAAAGCTTTAAGAGTTCCTTCAATTTCTCAT	1507
Qy	1643	CTAGAGATTGATCACCAAATTTACTTTGCCAATGCCAGCTATTATTAAGAGAAAGGATCACA	1702
Db	1508	TTGGCTGTGAGTCTCAATCTATTTTGCTTAATCACTTATCTTCAAGAAAGGATACTG	1567
Qy	1703	AGTGGATTGATGAAGAAGAAAGAAATTAAGCTACAGGGGAGACTAGTTTGCAGTAT	1762
Db	1568	AGATGGGTTGGAAGAGGAAGACATATAAAGCTTAATAATGGAGCTCCCATTTGAAGTGC	1627
Qy	1763	GTATATAATTGATGCTGCTGTTGGAACATTTGATACAGTGGATAGTATGCTTGAA	1822
Db	1628	ATAATTTTAGACATGACAGCTGTGCACAGCCACAGACACAAGTGGCTTGACACTTTATGT	1687
Qy	1823	GAGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTTGTGTTGGTCAATCTGTAAGT	1882
Db	1688	GAATTTAGAAGATGCTGGAGAGAGATCACTTGAAGTTTGTGCTGGCAATCTCTTTGGA	1747
Qy	1883	GAAGTGAAGAAGAACTGAACAAATCGAAGTTCACAAATCAATTTAGGAAGAAATGGATC	1942
Db	1748	AATGTGATGAAAAAATTGCAATAGTCAAGTCAACATTTTGGATTCTTTTGGATTAAAAAGGAGTC	1807
Qy	1943	TATCTGACTGTTGAAGAGCGCT	1965
Db	1808	TATCTCAGTGGGGAAGCTGT	1830

RESULT 13	
AJ704374	
LOCUS	AJ704374 2299 bp mRNA linear PLN 10-MAY-2004
DEFINITION	Brassica oleracea var. acephala mRNA for plasma membrane sulphate transporter (ST3.4 gene).
ACCESSION	AJ704374
VERSION	AJ704374.1 GI:47109355
KEYWORDS	plasma membrane sulphate transporter; ST3.4 gene.
SOURCE	Brassica oleracea var. acephala (kale)
ORGANISM	Brassica oleracea var. acephala Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE	1
AUTHORS	Buchner,P., Stuiver,E.E., Hawkesford,M.J. and de Kok,L.J.
TITLE	Analysis of the effect of sulphur nutrition and atmospheric H2S exposure on the expression of sulphate transporter genes in Curley cale
JOURNAL	Unpublished
AUTHORS	2 (bases 1 to 2299)
TITLE	Buchner,P.
JOURNAL	Direct Submission

JOURNAL	Submitted (06-MAY-2004) Buchner P., Crop Performance Improvement, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UNITED KINGDOM
FEATURES	Location/Qualifiers
source	1..2299
gene	/organism="Brassica oleracea var. acephala"
5' UTR	/mol_type="mRNA"
	/variety="acephala"
	/db_xref="taxon:3713"
	1..2299
	/gene="ST3.4"
	1..138
	/gene="ST3.4"
	/evidence=experimental
	139..2109
	/gene="ST3.4"
	/codon_start=1
	/product="plasma membrane sulphate transporter"
	/protein_id="CAG28416.1"
	/db_xref="GI:47109356"
	/translation="MGHGTNRMEDMSSPNNETAANARETVVEIHSVCLPPKTTFOKL
	KRFGDVPDPDPLERFRNQTRWKVILGLQSLFPIFPWGSQYDLKLFPSDVLISGLAI
	ASLAIPIQGISYAKLANLPPIVGLYSFVPLIYSLVSGSKHLAVGVSIASLVMSML
	SSVSPTQDSILYKLAPFTSTFAGLQSLRLRGLFAIDLSKATLVGFTAGAAYI
	VSLOQKDLGIYHFTGMQFVPVMSVINTSRSEWETIVMGLFLLIILVTRHISM
	RKPKLFWISAASPLASVISTLVVIRDKTHAIFIGHLPKGLNPPSNMILYFSAAH
	LALAKTGIITGLISVITGIAVGRFASLKNYQVNGKEMMAIGPMNMGASCTSCVVT
	TGSPRSVNVNAGAKTAVSNVIMASAVLCTLLFLMPLFYPTPNLLIAIILTAVIGL
	IDYPAIKLWKVDKDFDFTCMGSPFFGVLPVAVPLGAIAGVSVIKILHLVTRPNTLE
	FNQITQIYQSLKRYREASRIPGLIILAVESPIYFANCTYQERISRWTRREENRIK
	ENNERNLKCIILDMTAVSIDSIGIESVPELRRLRLENSQLVLNVPVSGVMEKLHKS
	KIISLGLSLGLYLTVGEAVSDLSSTWKAHQP"
	2110..2299
3' UTR	/gene="ST3.4"
	/evidence=experimental
ORIGIN	
Query Match	27.8%; Score 681; DB 8; Length 2299;
Best Local Similarity	61.0%; Pred. No. 3.8e-158;
Matches 1104; Conservative	0; Mismatches 705; Indels 0; Gaps 0;
Qy	158 TTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACTTCTTCCCTGATGACCTTTGAGGCAG 217
Db	259 TTTCAAAACTCAAAAAGCGTTTGGTGATGTTTCTTCCCTGATGATCCTTTAGAGAG 318
Qy	218 TTCAAGAACAGCCAGCTTCCAGAGAGTTCATGCTTGGCTTCAGTTCCTTCCCATTT 277
Db	319 TTAGAACACGACATGGAGAAACAAAGTCATTCTTGGTCTACAAAGCTTGTTCCTATA 378
Qy	278 TTGGAATGGCTCCCAAAATACACCTTTTCAGTTCTTTGAAAGCTGACCTCATAGCTGGCATC 337
Db	379 TTCCCATGGGTTCTCAGTACGATCTCAAGCTTTTATAGTCTGATGTTATCTGGACTC 438
Qy	338 ACCATGCTGATCTGGCCATTCCTCAGGCAATCAGTTATGCCAAGCTGCCAACCTCCCT 397
Db	439 GCCATTGCCAGTCTCGCTATCCCTCAGGGAATCAGTTATGCCAAGCTAGCAAACTTACCA 498
Qy	398 CCAATTCCTGGACTATATTCGAGCTTTATACACCAATGATTATGCGNATGATGGGTAGC 457
Db	499 CCCATTGTTGGTCTTTATTCAAGCTTCGTGCCGCCCATTTATATCTCAGTTCCTGGAGT 558
Qy	458 TCGAGGATTTTGGCAGTGGGAGCTGTGGCGGTGGATCGCTTCTGATGGGTTTCGATGTTG 517
Db	559 TCTAAGCATTTAGCAGTGGTCTCTCTATAGCATCACTTGTATGGGCTCAATGCTA 618
Qy	518 AGTAATGCGTTGATCCCAATGAAGACCCAAAGCTTTTACCTCCACCTGGCTTTCACAGCT 577
Db	619 AGTGAAGAGTGTCTTCTCAACACAGACTCAATATATATCTCAAACTGGCTTTCACATCA 678
Qy	578 ACATTATTTGCTGGTGTGTTTTCAGAGCTGCTTGGGCTGCTTTAGGTTGGGTTGATCGG 637
Db	679 ACTTTCTTGTGGCCTCTTCCAAAGCTCCCTTGGCCTTCTCAGGCTGGGATTTGGGATT 738

Qy	638	GATTTTCTGTGCATGCAACCATAAATAGGGTTCATGGGAGGAGCAGCCACGGTGGTGTGT	697
Db	739	GACTTTTTTGTCAAAGGGCAACATTTGGTTGGTTTCACTGCTGGTCTGCAGTAATTTGTGTCA	798
Qy	698	CTGCAGCAACTAAAATCGATTCTTTGGCTTGCAGCATTTCACCCATGAGCTGATATCATTA	757
Db	799	CTTCAACACAGCTTAAGGGTCTTTCTTGGAAATCGTTCACTTCACTGSCAAAAATGCAATTCGTT	858
Qy	758	TCAGTGTAGCGCTCTGTGTTTTCAACCACAACTCATGAGTGGAGTGGGAAAGTCTGTCTGTTA	817
Db	859	CCTGTCATGCTCTGTGTTATCAACACTAGATCCGAATGGTCATGGGAAACTATTCGTGATG	918
Qy	818	GGATGTGTCTCAATTTTCTTCCCTCTTAGCAACAAGATCTTTCAGCAAAAAACGACCAAGG	877
Db	919	GGCCCTTGGCTTCTTTGATCAATCTCTTAACCAACAAGACACATTAGCATGAGGAGGCCAAAA	978
Qy	878	TTTTTTTTGGGTGCAGCAATGGCGCCATTCAGCTCCGTTATATTGGGAAGTCTCTTGGTT	937
Db	979	CTTTTCTGGGATATCAGCTGCATCACCATTGGCATTCAGTTGTTAATCTCAACTCTTCTTGTGA	1038
Qy	938	TATTTCTACTCAGCGCGGAGAACCGGTGTTTGAAGTGATAGAGAACTGAAGAAGGGTTTG	997
Db	1039	TACGTATCAGAGACAAAACTCATGCCATCTCTTTTATTGGCCATCTGCCAAGGGTTTG	1098
Qy	998	AATCCACATCATCTCACAAAATCTGGTATTTGTGTGCGCTTACATGACTACAGCTGTCAAA	1057
Db	1099	AATCCTCCTTCAGCGAATATGTTGTACTTTAGTCTGCTCATCTTGTCTTTGCCATCAAG	1158
Qy	1058	ACTGGCAATTCGTTGGCATCATATCACTTTCGCGAAGGAATAGCAGTAGAAGAAGCTTT	1117
Db	1159	ACTGGTATCATCAGAGGATTTCTCTCCCTTACAGAAGGGATGCTGTAGGAGAACCTTT	1218
Qy	1118	GCAATGATATAAAATTAACAATATTGATGGCAACAAAGAGATGATAGTCTATTGGGACCATG	1177
Db	1219	GCATCTCTAAGAACTATCAAGTTAATGGGAAACAAAGAATATGATGGCTATAGGTTTTATG	1278
Qy	1178	AACGTAGTTGGTTCTTTTCACTCTTGCTACTCTCAACACAGGACCATTTTCGCGTTCGGCT	1237
Db	1279	AACATGCTGGTTCTTGACCTCTTGCTATGTTTACAACAGGGTCTTTTCTCAAGATCTGCA	1338
Qy	1238	GTGNACTATAACGCTGTGATGCAAGACAGCAGCTTCCAAACATTAATATGTCATCTTGACGTA	1297
Db	1339	GTAACGCTCAATGCTGGAGCGAAACAGCAGTTTCAAACATATGTGATGGCATCAGCAGTT	1398
Qy	1298	ATGTTGACATTGTTATTTCCTGCACACCCTGTGTTCCATTACACTCCCTCGTGGTGTCTATCA	1357
Db	1399	CTTGGGACCCCTTTGTTTTGATGCCACTCTTCTATACACTCTTAATCTGATCTCTAGCC	1458
Qy	1358	GCTATTATCGTATCTGCAATCTCTTGGACTCATAGATTATGAAGCAGCCATCCATCTATTT	1417
Db	1459	GCCATCATTTTAAACCGCTGTGATAGGCCCTCATTTGATTTACCAACGGCTTACAAGCTTTGG	1518
Qy	1418	AAGGTTGACAAATTTGACTTTTGGTGTGCATGAGTGCATACATTTGGCGTGGTCTTTGGC	1477
Db	1519	AAAAGTTGACAAATTCGATTTTCTTCACCTTGCAATGTGTTTCAATCTTTTGGCGTCTCTTTGTC	1578
Qy	1478	AGTGTGTAATTTGGCTTAGTCATAGCTATTGTAATATCTGACTTCGGGTACTTCTATTT	1537
Db	1579	TCGTGCTCTTTGGCTTAGCAATAGCAGTGGGAGTTTCAGTTATCAAGATCTTTGTTGCAT	1638
Qy	1538	ATTGCAAGGCCAAGGACATTCGTTTTGGGCAACATTCCAAATTCGTGTGATATACCGAAAT	1597
Db	1639	GTCAACAAGCCAAACACTTTTAGAGTTTCGGAACATCCCTGAGACTCAGATATATCAGAGT	1698
Qy	1598	GTTGAGCACTATCAAAATGCAAAACATGTTTCTCGGAATGCTAATTTCTAGAGATTGATGCA	1657
Db	1699	CTTAAAAAGATACAGAGAAGCCTCTAGAATCCCTCGTTTCTTTGATTTCTTGCAGTTGAATCT	1758
Qy	1658	CCAATTTACTTTGTCCAATGCCAGCTATTTAAGAGAAGGATCACAGGTCGATTGATGTA	1717
Db	1759	CCAAATACTTTTGCCAACTGCATTTACCTTACAAGAAGGATCTCGAGATGGAACGAGGAA	1818
Qy	1718	GAAGAAGAAAGAAATTAAGCTACAGGGGGAGACTAGTTTGCAGTATGTTATTAATTGATATG	1777

Db 1819 GAGGAAAAATCGGATCAAGGAAAAACAATGAGAGAAACCTTAAAGTGCAATTATCTTGCATG 1878
 Qy 1778 AGTCTCTTGGAAACATTTGATACAGTGGAAATAGATAGTCTTGAAGAGGTGAAGAGATT 1837
 Db 1879 ACAGCTGTGCTTCCATAGACACAGTGGGATGGAATCGGTGTTGAACATTAGGAGGAGG 1938
 Qy 1838 ACAGAGAGAGAGAGCTACAGCTGTGTTTGGTCAATCTCTGAAGTGAAGTGAAGAGAA 1897
 Db 1939 CTGAGAAATCAATCACTTACGCTTGCTGCTGTAATCTCTGGGAGTGTATGGAAG 1998
 Qy 1998 CTGAACAAATCGAAGTTCCAAATCAATTTAGGGAAGAAATGGATCTATCTGACTGTGAA 1957
 Db 1999 CTACACAAGTCCAAGATCAATGAGTCAATGGGCTGAGTGGGCTTATCTTAAACAGTTGGT 2058
 Qy 1958 GAGGCCGTT 1966
 Db 2059 GAAGCTGTT 2067

RESULT 14
 AB054645 Arabidopsis thaliana 2256 bp mRNA linear PLN 23-JAN-2001
 LOCUS Arabidopsis thaliana mRNA for sulfate transporter Sultr3;4,
 DEFINITION complete cds.
 ACCESSION AB054645
 VERSION AB054645.1 GI:12381948
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (sites)
 Takanashi, H., Watanabe-Takahashi, A., Saito, K. and Yamaya, T.
 cDNA for sulfate transporter Sultr3;4
 Published Only in DataBase (2001)
 2 (bases 1 to 2256)
 Takanashi, H., Watanabe-Takahashi, A., Saito, K. and Yamaya, T.
 Direct Submission
 Submitted (19-JAN-2001) Hideki Takahashi, RIKEN Plant Science
 Center; Hirosewa 2-1, Wako, Saitama 351-0198, Japan
 (E-mail:hideki@postman.riken.go.jp, Tel:81-48-467-6899,
 Fax:81-48-467-6857)

FEATURES
 Location/Qualifiers
 1..2256
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 95..2056
 /codon_start=1
 /product="sulfate transporter Sultr3;4"
 /protein_id="BAB21264.1"
 /db_xref="GI:12381949"
 /translation="MGHGNTNRVEDMASPNNGTAGETVWEIHSVCLPDKKTAFOKKLR
 VGVDFPDPLQRFNPTQWRNRVILGQLSPFTFTWGSVDLKLRSDDVTSGLTIALS
 AIPQGISYAKLANLPIVGLVSFVPLIYAVLGSRHLAVGPVSLAVMGSMLES
 VSTQBSILYKLAFTSTFPAGVFQASGLRLGFWIDFSLKALCTGFTAGAAVSL
 QVQQLGVHVFTHGRNQIVPVMSSVFNHRSEWSTVIMVGIGLSTLLTRHISMKPK
 KLFWISAGTPLASVITISLLVLRISKTHAIPFHLKPLNPPSLNMLYFSGAHIAL
 ALTIAGITGLSITEGIAVGRTPASLKNQVNGNKEMMAIGFNMAGSCTSCVVTGS
 FRSAYVYNAGATCAVNTVMASAVLTLFLMPLPYTTPNVIILAIILTAVTGLDY
 QAAKLYKWDKDPFPCLSGFVSPVGLAJAVAVSVIKILLHVRPNPTSEGN
 IPTQYIKSLGRYREASIPGFHLLAIESPIYFANSTYQDRLRWARBEENRIKKN
 GTTLKCIILDMTAVSAIDTSGLEAVFELRRLEKOSLQLVLVNPVGTVMVKLHKSKI
 EALGLSLYLTVGEAVADLTSSTWKANGQP"

ORIGIN
 Query Match 27.7%; Score 678.6; DB 8; Length 2256;
 Best Local Similarity 60.5%; Pred. No. 1.5e-157;
 Matches 1116; Conservative 0; Mismatches 722; Indels 0; Gaps 0;
 Ov 123 TGACCAAGTCGAGGTTCCACGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGA 192

[illegible]

Db	1251	GCTATGTTACAA	CAGGGTCTTTCTCTCGATCTCGGGTAAACATATAATGCTGGAGCGAAA	1311
Qy	1263	CAGCAGCTTCCAA	CATTATAAATGTCATCTTGCGAGTAATGTTGACATATGTTATTCCTGACAC	1322
Db	1311	CAGGGTTTCTACA	TGTGATGGCCTCTGCAGTCTTAGTGACCCCTCTGTTTTGATGC	1370
Qy	1323	CTTTGTTCCAA	TACATCCCTGGTGGTCTATCAGCTATTAATTCGTAATCGTAATGCTTG	1382
Db	1371	CACCTCTTCTATT	ACACTCTTAATGTGATCCTAGCTGCTATCATCTTTAAACCGCGGTGATAG	1430
Qy	1383	GACTCATAGAT	TATGAAGCAGCCATCCATCTATTTAAGGTTTGACAAATTTGACTTTGTGG	1442
Db	1431	GCCTCATGT	ATATCAAGCGGCTTCAAGCTTTTGGAAAGTTGACAAATTCGATTTCTTCA	1490
Qy	1443	TGTGCATGAGT	GATACATTTGGCGTGGTCTTTTGGCAGTGTGAAATTTGGCTTAGTCATAG	1502
Db	1491	CTTTGCTTGTG	TTCTTTTGGGCTTCTCTTTGTTCTTGTAACCTTTGGCCTAGCAATAG	1550
Qy	1503	CTATTGTAAT	CTGTACTTTCGGGTACTTCTATTTATTCGAAGCCAAAGACAATTCGTTT	1562
Db	1551	CAGTGGCAG	TTTCAGTATCAAGATCTGTGTGCACGTAAACCGAGCCAAACACTTCAGAAT	1610
Qy	1563	TGGGCAACAT	TCCAAATTTCTGTGATATACCGAATGTTGAGCACTATCAAAATGCAAAAC	1622
Db	1611	TTGGGAAT	TCCAGGGACTCAGATTTTATCAGAGTCTTGGAAAGTACAGAGAAGCTTCAA	1670
Qy	1623	ATGTTCTGGA	ATGCTAAATTTCTAGAGATGATGCACCAATTTACTTTTGGCAATGCCAGCT	1682
Db	1671	GAATCCAGG	CTTCTTAATTCCTTATTTGAATCTCTATATATTTTCGCNATAGCACTT	1730
Qy	1683	ATTTAAGAGA	AAAGGATCACAAGTGGATGATGAAGAAGAAGAAATTTAAAGCTACAG	1742
Db	1731	ATCTTCAAG	ACAGCATTTTGGAGATGGCCAGGGAAGAGGAAAAATCGATAAAGGAGAAC	1790
Qy	1743	GGAGACTAG	TTTGGAGTATGTTATTAATTTGATATAGTCTGTTTGGAAACATTTGATACAA	1802
Db	1791	ATGGTACT	ACTTCTGAAATGTCATTAATTTCTTGACATGACAGCTGTCTCGCGATAGACAA	1850
Qy	1803	GTGGAATA	GATATGCTTTGAAGAGGTGAAGAAGATTTACAGAGAAGAGAGCTACAGCTTG	1862
Db	1851	GCGGGCTT	GAGCGGTGTTTGAACTTAGAGGAGACTCGAGAAGCAATCTCTTCAGCTTG	1910
Qy	1863	TTTTTGGT	CAATCTCTGAAGTGAAGTGAAGAAACCTGAACAAATCGAAGTTCCAAAATC	1922
Db	1911	TGCTGTGTA	ATCTCTGTGGGACTGTGATGAAAAGCTACACAAGTCCAAGATCATTTGAGG	1970
Qy	1923	ATTTAGGGA	AGAATGGAATCTATCTGACTGTTTGAAGCGCGTTG	1967
Db	1971	CATTGGGT	CTGAGTGACATTTATCTAACAGTTTGTGTAAGCTGTGG	2015
RESULT 15				
AF355602				
LOCUS				
DEFINITION				
Zea mays sulfate transporter ST1 mRNA, complete cds.				
ACCESSION				
AF355602				
VERSION				
AF355602.1				
KEYWORDS				
GI:13625940				
SOURCE				
Zea mays				
ORGANISM				
Zea mays				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD				
clade; Panicoideae; Andropogoneae; Zea.				
1 (bases 1 to 2640)				
Hopkins,J., Parmar,S., Bouranis,D.L., Howarth,J.R. and				
Hawkesford,M.J.				
Coordinated Expression of Sulfate Uptake and Components of the				
Sulfate Assimilatory Pathway in Maize				
Plant Biol. 6 (4), 408-414 (2004)				
2 (bases 1 to 2640)				
Hopkins,J., Bouranis,D.L. and Hawkesford,M.J.				
Full length cDNA of a high-affinity type sulfate transporter from				
Zea mays				

JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 2640)
AUTHORS	Hopkins,L., Bouranis,D.L. and Hawkesford,M.J.
TITLE	Direct Submission
JOURNAL	Submitted (01-MAR-2001) Agriculture and Environment Division,
FEATURES	IACR-Rothamsted, Harpenden, Hertfordshire AL5 2JU, UK
source	Location/Qualifiers
	1..2640
	/organism="Zea mays"
	/mol_type="mRNA"
	/db_xref="taxon:4577"
CDS	120..2096
	/codon_start=1
	/product="sulfate transporter ST1"
	/protein_id="AAK35215.1"
	/db_xref="GI:13625941"
	/translation="MPERTVSDGDDNGTDVASRTSSHRETEHGHYKVGVPPKKNL AETSDAVKETFPFADDPRLQYKDQRSKKIWLGHQIFPVLDNSRHSLGKFPGKDFATAG TLTASUCIPDITGGSEKLPAEVLGYSSFPPLIYAIVGSSRDIAIGPVAVVSVLLILG TLINACEDPHGTLEVRALPTAFVFAGVTQAALGFPRLFGIIIEFLSHAAIVGFMAGA AITIALQLQGFLGICIANFTKKSIDVISVMKSWGNVHGMNWTILIGATFLAPLVAK YIKGRKNLKFWSAIAPLTVISITPFVITRADKHGVAIVKNIRGINPPPSASLIYV TGPLYLATGFKIGIVAGMIGEITAIAIGRTPAALKDYRIDGNKEMWALGTNNIVGSITS CVYATGFSRSAYSNMAGCTAVSNVVMISLMLLLLILPLPKTFPNALLSIISIILA VLGLDIYESAILWKVDKDLFLACMGAFVGIVFESSVEYGILLIAVALSKLILLOVTRP EVOODOKLTKEFLVIDLSVIVIDISTGIIHALBELAKLEKRKIQLVLTNPGPVAVIQK LRSAKFTMDIMGEDNIFLTVGDVAKVPKMNVDV"
ORIGIN	
Query Match	27.7%; Score 677.4; DB 8; Length 2640;
Best Local Similarity	60.8%; Pred. No. 3e-157;
Matches 1123; Conservative	0; Mismatches 721; Indels 3; Gaps 1;
Qy	122 GTGCACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTG 181
Db	213 GGGTACAAGATCGCGCGTCCCACCGGAAGAACCTCTCCTCCCGAGATCTCTGACGGGTG 272
Qy	182 AAGGAGACTTTCTTCCCTGTATGACCCCTTTGAGGCAGTTCAAGAACAGCAGCTTCCAAG 241
Db	273 AAGGAGACGTTCTTCGCCGACGACCCGCTGGGCAGTACAAGNACAGCCAGGTCUAG 332
Qy	242 AAGTTCATGCTTG3CCTTCAGTTTCTTCTTCCCCATTTTTCGAATGGGCTCCAAATACACC 301
Db	333 AAGATCTGGCTGG3CCTCGACGACATCTTCCCGGTGCTCGACTGGACGAGCATTACTCC 392
Qy	302 TTTTCAGTTCTTGAAGCTGACCTCATAGCTGGGATCACATCGGTAGCTTGGCGCATTCCT 361
Db	393 CTGGCAAGTTCAGGGTGATCTTCATCGCTGGGCTCACCATTCGCAGCCTCTGCATACCC 452
Qy	362 CAGGGCATCAGTTATGCAAAGCTCGCCAACTCCCTCCAAATCTTGTGACATATATTGAGC 421
Db	453 CAGACATCGGTATTCTAAGTTTGCTTAACTTGCACAGAGGTGGGACATATACAGTAGC 512
Qy	422 TTTTATACCACTATTGATTATGCGATGATGGGTAGCTTCGAGGGAATTTGGCAGTGGGACT 481
Db	513 TTCTGTCGGCTCTGTATATACGCAGTGATGGCGAGCTCCAGGGACATAGCCATCGGCCCA 572
Qy	482 GTGGCGGTTCGATCGCTTCTGATGGGTTTCGATGTGAGTAATAACGGTTGATCCCAATGAA 541
Db	573 GTGGCCGTGTGTGCTGTGCTCGGTACTCTCCCTTCAGAATGAGATCGACCCCAAGACA 632
Qy	542 GACCCAAGCTTTTACCTCCACCTGGCTTTCACAGCTACATTTATTGCTGCTGTTTTTTCAG 601
Db	633 CACCCGCTCGAGTACAGCGGCTTAGCCCTTCACAGCAACCTTTTTTCGACGGGGTCACTCAG 692
Qy	602 GCTGCTTTGGGTCTGTGTTAGGTTTGGGTTTGATCGTGGATTTTCTGTGCATGCAATCCATA 661
Db	693 GCAGCCCTCGGATCTTCAGGCTAGGGTTTATCATAGAGTTCTTGTCTCATGTGCGCAATT 752
Qy	662 ATAGGGTTCAATGGAGAGAGCAGCCAGCGGTGGTGTCTGCAGCAATTAATAATCGATTCCT 721
Db	753 GTCGGATTATGCGCGCGCTGCGCATCACATTTGCGCTTCAGCAGCTTGAAGGATTCCTT 812

Qy	1799	ACAAAGTGGAAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAG	1858
Db	1893	ACAAAGTGGAAATCCATGCTTTTGGAGGAGTTGGCGAAAGCTCTTGAAAAACGCAAAATTCAG	1952
Qy	1859	CTTGTGTTTGGTCAATCCTGTAAGTGAAGTGAAGAAACTGAACAAATCGAAGTTCCAA	1918
Db	1953	CTGGTTCTTACCAATCCCGGCGCGGGTGATCCAGAAGCTCCGCTCAGCGAAATTCAG	2012
Qy	1919	AATCATTTAGGGGAAGAAATGGATCTATCTGACTGTTGAAGAGGCCGT	1965
Db	2013	GACATGATTGGTGAAGACAACATATTCCTCAGGTCGGCGACGCTGT	2059

Search completed: August 31, 2005, 06:33:50
Job time : 10593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 19:18:05 ; Search time 1312 Seconds
(without alignments)
11049.883 Million cell updates/sec

Title: US-10-762-049-17
Perfect score: 2449
Sequence: 1 gcacgagctagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2449	100.0	2449	3 AAZ50489	Aaz50489 Soybean s
2	1041.2	42.5	1977	3 AAC43310	Aac43310 Arabidops
3	1038	42.4	2112	3 AAC43093	Aac43093 Arabidops
4	1018.2	41.6	1977	6 ABZ13697	Abz13697 Arabidops
5	1018.2	41.6	1977	8 ADA68514	Ada68514 Arabidops
6	846	34.5	1974	8 ADA70002	Ada70002 Rice gene
7	765.2	31.2	1981	3 AAZ50482	Aaz50482 Corn sul
8	720.6	29.4	2067	3 AAZ50488	Aaz50488 Soybean s
9	662.4	27.0	2006	3 AAC42229	Aac42229 Arabidops
10	660.2	27.0	1896	8 ADA68554	Ada68554 Arabidops
11	655.4	26.8	2279	3 AAZ50481	Aaz50481 Corn sul
12	650.4	26.6	1989	8 ADA70466	Ada70466 Rice gene
13	647.2	26.4	1947	8 ADA69770	Ada69770 Rice gene
14	640.4	26.1	2311	3 AAZ50490	Aaz50490 Wheat sul
15	634.4	25.9	1971	6 ABQ82719	Abq82719 Arabidops
16	630.4	25.7	1940	8 ADA71315	Ada71315 Rice gene
17	600.2	24.5	2000	8 ADA70110	Ada70110 Rice gene
18	559.8	22.9	1896	8 ADA69681	Ada69681 Rice gene
19	455.2	18.6	1983	8 ADA69804	Ada69804 Rice gene
20	401.4	16.4	1980	12 ADJ40066	Adj40066 Plant cDN

21	398.8	16.3	1974	8 ADA70351	Ada70351 Rice gene
22	324	13.2	780	3 AAZ50484	Aaz50484 Artichoke
23	217.2	8.9	2058	6 ABZ12912	Abz12912 Arabidops
24	217.2	8.9	2058	8 ADA68540	Ada68540 Arabidops
25	216.2	8.8	1240	3 AAZ50483	Aaz50483 Corn sul
26	210.8	8.6	500	10 ABX61505	Abx61505 Arabidops
27	208.8	8.5	490	10 AD881920	Ad881920 Arabidops
28	205.6	8.4	2022	3 AAZ50491	Aaz50491 Wheat sul
29	184.8	7.5	596	12 ADJ43778	Adj43778 Plant cDN
30	148.8	6.1	552	12 ADJ44485	Adj44485 Plant cDN
31	129.6	5.3	699	9 ACL17188	ACL17188 DNA clone
32	124.6	5.1	695	9 ACL17186	ACL17186 DNA clone
33	123.2	5.0	680	9 ACL17191	ACL17191 DNA clone
34	118	4.8	493	3 AAZ50487	Aaz50487 Rice sul
35	117.4	4.8	484	3 AAZ50485	Aaz50485 Rice sul
36	115.2	4.7	658	9 ACL17184	ACL17184 DNA clone
37	104.4	4.3	630	9 ACL17187	ACL17187 DNA clone
38	100.2	4.1	264	6 ABL74697	Ab174697 Corn tass
39	97.4	4.0	2241	3 AAZ29466	Aaz29466 P. chryso
40	94.8	3.9	594	9 ACL17185	ACL17185 DNA clone
41	93.4	3.8	581	9 ACL17189	ACL17189 DNA clone
42	90.4	3.7	2441	11 ADM57788	Adm57788 Murine pr
43	90.4	3.7	2441	12 ADE86099	Ade86099 cDNA enco
44	89.4	3.7	2459	6 ABQ59208	Abq59208 Listeria
45	89.4	3.7	2692	6 ABQ70888	Abq70888 Listeria

ALIGNMENTS

RESULT 1
AAZ50489
ID AAZ50489 standard; cDNA; 2449 BP.
XX
AC AAZ50489;
XX
DT 23-MAY-2000 (first entry)
XX
DE Soybean sulphate permease-2 cDNA clone.
XX
KW Sulphate Permease; sulphate assimilation protein; soybean; probe;
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
KW antibody; screen; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 2..2044
FT /*tag= a
FT /product= "Soybean sulphate permease-2"
FT /note= "Derived from clone sfill.pk0043.g10"
XX
PN WO200004154-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015810.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Falco SC, Thorpe CJ;
XX
DR WPI: 2000-195025/17.
DR P-PSDB; AAY44943.
XX
PT Nucleic acid fragments encoding sulfate assimilation proteins in plants
PT and seeds useful as probes for isolating cDNAs and genes encoding
PT homologous proteins, in producing transgenic plants.
XX
PS Claim 3; Page 52-53; 79pp; English.
XX

CC The present sequence is a cDNA clone encoding soybean sulphate permease,
CC a sulphate assimilation protein. This sequence is obtained from sf11
CC library, clone sf11.pk0043.g10, derived from soybean immature flower.
CC This sequence is used as a probe to isolate other plant sulphate
CC assimilation proteins, for genetic and physical mapping of related genes
CC and as markers of traits linked to the gene. This is useful for plant
CC breeding and to construct chimeric genes, used to create transgenic
CC plants with altered levels of sulphate permease. The sulphate permease
CC peptides are useful for producing antibodies, that are used to screen and
CC isolate cDNA clones
XX
SQ

Sequence 2449 BP; 719 A; 470 C; 543 G; 717 T; 0 U; 0 Other;

Query Match 100.0%; Score 2449; DB 3; Length 2449;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GCACGAGCTAGCTCGCACATTAAGTTATATATAACACATATATTTGCTTGTAGAAATACTAT	60
Db	1	GCACGAGCTAGCTCGCACATTAAGTTATATATAACACATATATTTGCTTGTAGAAATACTAT	60
Qy	61	TATTGAAGATATGGGAGTGTAGATTATATAGTACCCCTTTGGGCATGAACAACTTTGAGAG	120
Db	61	TATTGAAGATATGGGAGTGTAGATTATATAGTACCCCTTTGGGCATGAACAACTTTGAGAG	120
Qy	121	AGTGCACCAAGTCGAGGTTCCACCGCCACACGCGCTTTTCAAGTCTCTAAAGTACTCTTT	180
Db	121	AGTGCACCAAGTCGAGGTTCCACCGCCACACGCGCTTTTCAAGTCTCTAAAGTACTCTTT	180
Qy	181	GAAGGAGACTTCTTCCCTGATGACCCCTTTGAGGCGAGTTCAGAAACAAGCAGCTTCCAA	240
Db	181	GAAGGAGACTTCTTCCCTGATGACCCCTTTGAGGCGAGTTCAGAAACAAGCAGCTTCCAA	240
Qy	241	GAAGTTTCACTTTGGCTTCAGTTCCTTCCCAATTTTCGAATGGCTCCCAAAATACAC	300
Db	241	GAAGTTTCACTTTGGCTTCAGTTCCTTCCCAATTTTCGAATGGCTCCCAAAATACAC	300
Qy	301	CTTTTCAGTCTTGAAGAGTCGACCTCATAGTGGCATCACCATCGCTAGCTTGGCCATTC	360
Db	301	CTTTTCAGTCTTGAAGAGTCGACCTCATAGTGGCATCACCATCGCTAGCTTGGCCATTC	360
Qy	361	TCAGGGGATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATTCCTTGGAGTATATTCGAG	420
Db	361	TCAGGGGATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATTCCTTGGAGTATATTCGAG	420
Qy	421	CTTTATACCAACCATGATTTATCGGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAC	480
Db	421	CTTTATACCAACCATGATTTATCGGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAC	480
Qy	481	TGTGGCGTTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTTGATCCCAATGA	540
Db	481	TGTGGCGTTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTTGATCCCAATGA	540
Qy	541	AGACCCAAAGCTTTACCTCACCTGGCTTTCACAGCTACATTAATTTGCTGGGTGTTTTCA	600
Db	541	AGACCCAAAGCTTTTACCTCACCTGGCTTTCACAGCTACATTAATTTGCTGGGTGTTTTCA	600
Qy	601	GGCTGCTGGGTCTGTTTATAGGTTGGGTTGATCGTGGATTTTCTGTCACTGCAACCAT	660
Db	601	GGCTGCTGGGTCTGTTTATAGGTTGGGTTGATCGTGGATTTTCTGTCACTGCAACCAT	660
Qy	661	AATAGGGTTTCATGGGAGGAGCCACCGTGGTGTGTCTGCGCAAACTAAAATCGATTCT	720
Db	661	AATAGGGTTTCATGGGAGGAGCCACCGTGGTGTGTCTGCGCAAACTAAAATCGATTCT	720
Qy	721	TGGCTTGTAGCATTTACCCATGGAGCTGATATCATATCATGATGCGCTCTGTTTTAC	780
Db	721	TGGCTTGTAGCATTTACCCATGGAGCTGATATCATATCATGATGCGCTCTGTTTTAC	780
Qy	781	CCAAACTCATGAGTGGAGGTGGAAAGTGTGTTGAGGATGTCTTCAATTTCTTCTCT	840
Db	781	CCAAACTCATGAGTGGAGGTGGAAAGTGTGTTGAGGATGTCTTCAATTTCTTCTCT	840

Qy	841	CCTTAGCACAAGATACCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTCTAGCAATGCG	900
Db	841	CCTTAGCACAAGATACCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTCTAGCAATGCG	900
Qy	901	GCATTTGAGCTCGGTTATATTTGGGAAGTCTCTTGGTTTATTTTCTACACCGCGAGAAGCA	960
Db	901	GCATTTGAGCTCGGTTATATTTGGGAAGTCTCTTGGTTTATTTTCTACACCGCGAGAAGCA	960
Qy	961	CGGTGTTGAAGTGTAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATCT	1020
Db	961	CGGTGTTGAAGTGTAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATCT	1020
Qy	1021	GGTATTTGTCTGCTTACATCACTACAGCTGTCAAACTGGCATTTGCTTGGCATCAT	1080
Db	1021	GGTATTTGTCTGCTTACATCACTACAGCTGTCAAACTGGCATTTGCTTGGCATCAT	1080
Qy	1081	ATCCTTTGGGAAGGAATAGCAGTAGGAGAGCTTTGCAATGTATAAAAAATACAAATAT	1140
Db	1081	ATCCTTTGGGAAGGAATAGCAGTAGGAGAGCTTTGCAATGTATAAAAAATACAAATAT	1140
Qy	1141	TGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTACCTTC	1200
Db	1141	TGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTACCTTC	1200
Qy	1201	TTGCTACCTTCACAAACAGGACCAATTTTCGGCTTCGGCTGTGAACATATAACGCTGGATCAA	1260
Db	1201	TTGCTACCTTCACAAACAGGACCAATTTTCGGCTTCGGCTGTGAACATATAACGCTGGATCAA	1260
Qy	1261	GACAGAGCTTCCAAATTAATTAATGTCACTTCGAGTAATGTTGACATTTGTTTCTCTGAC	1320
Db	1261	GACAGAGCTTCCAAATTAATTAATGTCACTTCGAGTAATGTTGACATTTGTTTCTCTGAC	1320
Qy	1321	ACCTTGTTCATTAACATCACTCCCTCGTGTGTCTATCAGCTATTATTCGATATTCGCAATGCT	1380
Db	1321	ACCTTGTTCATTAACATCACTCCCTCGTGTGTCTATCAGCTATTATTCGATATTCGCAATGCT	1380
Qy	1381	TGGACTCATAGATTAATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTCT	1440
Db	1381	TGGACTCATAGATTAATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTCT	1440
Qy	1441	GGTGTGCATGAGTGATACATTTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAT	1500
Db	1441	GGTGTGCATGAGTGATACATTTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAT	1500
Qy	1501	AGCTATTGTAATCTGTACTTTCGGGTACTTCTATTTATTGCAAGGCCAAGGACATTCCT	1560
Db	1501	AGCTATTGTAATCTGTACTTTCGGGTACTTCTATTTATTGCAAGGCCAAGGACATTCCT	1560
Qy	1561	TTTGGGCAACATTCGAAATTCGATATACCGAAATGTTGAGCACTATCAAAATGCAAA	1620
Db	1561	TTTGGGCAACATTCGAAATTCGATATACCGAAATGTTGAGCACTATCAAAATGCAAA	1620
Qy	1621	ACATGTTCTCGAATGCTAAATTCAGAGATTGATGCACCAATTTTCTTTGGCAATGCCAG	1680
Db	1621	ACATGTTCTCGAATGCTAAATTCAGAGATTGATGCACCAATTTTCTTTGGCAATGCCAG	1680
Qy	1681	CTATTTAAGAGAAAGGATCACAAAGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTAC	1740
Db	1681	CTATTTAAGAGAAAGGATCACAAAGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTAC	1740
Qy	1741	AGGGGAGACTAGTTTGCAGATGTTTATTAATGATAGTGTCTGTGGAAACATTTGATAC	1800
Db	1741	AGGGGAGACTAGTTTGCAGATGTTTATTAATGATAGTGTCTGTGGAAACATTTGATAC	1800
Qy	1801	AAAGTGGAAATGATGCTTTGAAGAGGTGAAGAGATTACAGAGAGAGAGCTACAGCT	1860
Db	1801	AAAGTGGAAATGATGCTTTGAAGAGGTGAAGAGATTACAGAGAGAGAGCTACAGCT	1860
Qy	1861	TGTTTTTGGTCAATCTCTGTAAGTGAATGAAGAACTGAAACAAATCGAAGTTCCT	1920
Db	1861	TGTTTTTGGTCAATCTCTGTAAGTGAATGAAGAACTGAAACAAATCGAAGTTCCT	1920
Qy	1921	TCAATTTAGGGAAGAAATGGATCTATCTGACTGTTTGAAGAGGCCGTTGGAGCATGCAACTT	1980

[illegible]

Db	253	GAAACTCTGTTTCAGACGACCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAA	312
Qy	245	TTCAATGCTTGCGCTTCAGTTCCTCTCCCATTTTCGAATGGGCTCCCAATACACCTTT	304
Db	313	TTTGTTAGGCTCAAACTACTCTCCGATTTTCGAATGGGACCAACGCTACATCTC	372
Qy	305	CAGTTCCTGAAAGTGAACCTCATAGCTGGGATCAACCAATCGCTAGCTTGGCATTCCTCAG	364
Db	373	AAGTTCCTTCAAAATCAGATCTCATCGCGGAATCACCATCGCTAGCTCTCGCATCCCTCAG	432
Qy	365	GGCATCAGTTATGCAACAGCTCGCAACCTCCCTCCAAATCTTGACATATATTGAGGTTT	424
Db	433	GGCATCAGTTACGCCAAATTTGCTAACTTGCCCCCAATTTCTGGCCCTTATTTGAGTTT	492
Qy	425	ATACCAACATTTGATTTATGGCATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTGTG	484
Db	493	GTACCGCCATTTGATACGCGGTGCTAGGGAGTTCAAGGACTTAGCGGTGGGAACGGTT	552
Qy	485	CGGTTGGATCGCTTCTGATGGGTTGATGTTGATGTAATGCGGTTGATCCCAATGAAGAC	544
Db	553	CGGTTGCGTCTCTGTTGACAGGTGCGATGCTGAGCAAGAAAGTTGATGCTGAGAAAGAT	612
Qy	545	CCAAAGCTTTACCTCCACCTGGCTTTCACAGCTACATTTTGTGCTGTTTTCAGGCT	604
Db	613	CCTAAGCTTTACCTTTCACCTTTGCTTTCACCGCCACTTTTTCGCGCGGTTCTCGAAGCC	672
Qy	605	GCCTTGGGCTCTGTTTAGGTTGGGTTGATCGTGATTTTCTGTACATGCAACCAATAA	664
Db	673	TCCTTTGGAATTTTCAGTTAGGTTTCATAGTGAATTTCTATCGCATGCAACGATAGTA	732
Qy	665	GGGTTCAATGGAGAGCAGCACGCTGGTGTCTGAGCAACTAAATCGATTCCTGGC	724
Db	733	GGATTCATGGGAGAGCAGCACGCTGGTGTCTGCAACAGCTTAAGGGTATATTCCGA	792
Qy	725	CTTGAGCATTTCAACCATGAGCTGATATCATATCAGTGATCGCTCTGTTTTCACCCAA	784
Db	793	CTTAAACATTTTCAGACTCTACCGATGTTATCTCTGTCAATGCTTCCGTTTCTCCCAA	852
Qy	785	ACTCATGAGTGGAGTGGGAAAGTGTGTGTAGGATGTGTCTTCAATTTTCTTCTCTCT	844
Db	853	ACTCAGAGTGGAGATGGGAAAGTGGGTTCTTGGATGTGGTTTCTATTTCTTCTCTC	912
Qy	845	AGCAAAAGATCTTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGGCGCA	904
Db	913	TCCACCAAGATTTTCAGCATCAAGAAACCAAAATCTTTTGGGTGGCGGATGGCTCCT	972
Qy	905	TTGACGTCCTTATTTGGGAGTCTCTTGGTTTATTTTCACTCAACGCGGAGAGCAGGT	964
Db	973	TTGACCTCAGTGATTTCTTGGAAAGTCTCTTGGTTTACTTCACTCAACGCTGAGAGACATGGT	1032
Qy	965	GTTGAAGTGATAGGAACTGAAGAAAGGTTTGAATCCACCATCACTCACAAATCTGGTA	1024
Db	1033	GTTCAAGTGATAGGGACTGGAAGAGGTTGAATCCACTCTCCGTTCTGATCTCATC	1092
Qy	1025	TTTGTTGCGCTTACATGATACAGCTGTCAAACTGGCATTTGTGTTGGCATCATATCA	1084
Db	1093	TTTACTTCCCTTACATGTCCACAGCTGTCAAAACTGGCCTCATCACTGGCATCATGTCT	1152
Qy	1085	CTTGCGGAAGGATAGCAGTAGGAGAGCTTTGCAATGTATAAAATACAAATTTGAT	1144
Db	1153	CTCGCTGAAGAGTAGCAGTGGGAGGAGTTTTCGATGTTCAGAACTACAAACATAGAC	1212
Qy	1145	GGCAACAAAGAGATGATAGCTATTGGGACCATCAACGATGTGGTTCTTTCACTCTTGC	1204
Db	1213	GGGACAAAGATGATAGCGTTTGGAAATGATGAAATCGTTGGTTCTTCACTCTGT	1272
Qy	1205	TACCTCAACAGGACCAATTTTCGCTTGGCTGTGAATATTAACGCTGGATCAAGACA	1264
Db	1273	TACCTCAACCGGACCAATTTTCAAGGTTCGCGAGTGAATACAAACGGGGTTGCAAGACC	1332
Qy	1265	GCAGCTTCCAAACATTAATGTCTACTTTCGAGTAATGTTGACATGTTATTCCTTGACACC	1324

Db	1333	GCATGTCCAAACATAGTAGTGGCGATTGGGTTATGTTTCACTCTCTTCTTCTCACACCG	1392
Qy	1325	TTGTTCCATTTACACTCCCTCGTGGTGTCTATCAGCTATTATCTGTATCTGCAATGCTTGA	1384
Db	1393	CTTTTTCACTACACACCACTCGTGGTCTCTCTCGCCATCATATATCCGCAATGCTCGGA	1452
Qy	1385	CTCATAGATTATGAAGAGGCATCCATCTATTATTAAGTTGACAAAATTTGACTTTTGGTG	1444
Db	1453	CTCATTTGACTATCAAGCTGCCATCCATCTCTGGAAGTTTGACAAGTTTCGACTTCTCTGTC	1512
Qy	1445	TGCATCAGTGCATACATTTGCGTGGTCTTTTGGCAGTGTGAAAATTTGCTTTAGTCATAGCT	1504
Db	1513	TGCATGAGCGCTTACGTTGGGTGCTATTTCGGCAGTGTAGAGATTGAGCTCGTGTAGCG	1572
Qy	1505	ATTGTAATATCTGTACTTTCGGGTACTTCTATTATTTTTCGAAGGCCAAGGACATTCGTTTG	1564
Db	1573	GTGGCGATATCTATAGCGAGTTTGTGCTGTCGAGGCCAAAACATCGCGTGAAG	1632
Qy	1565	GGCAACATTTCCAAATTTCTGTGATATACCGAAATTTTGGACACTATCAAAATGCAAAACAT	1624
Db	1633	GGAAACATATCCAAACAGCATGATCTATAGGAACACTGAGCAGTACCCATCATCAAGAACC	1692
Qy	1625	GTTCTCGGAATGCTAATCTTAGAGATTGATGCACCAATTTACTTTTGCATATGCCAGCTAT	1684
Db	1693	GTTCTCGGTATTCTCATCTTGGAGATTGATCTCCCATCTACTTTGCTAATGCCAGTTAC	1752
Qy	1685	TTAAGAGAAAGGATCACAAGGTGGATTGATGAAGAAAGAAAGAAATTTAAAGCTACAGGG	1744
Db	1753	TTGCGTGAAGAATCATAGGTGGATTGATGAAGAGGAAGAGAGAGTTTAAACAATCAGGA	1812
Qy	1745	GAGACTAGTTTGCAGTATGTTATTAATTGATATGAGTGTGTTGGAAACATTGTATACAGT	1804
Db	1813	GAGAGCAGCTTACAATATATTATCTGATTCGATATGTCAGCTGTTTGGTAAATATCGACACAAGC	1872
Qy	1805	GGAAATAGTATGCTTCAAGAGGTGGAAGAGATTACAGAGAGAGAGAGCTTACAGCTTGT	1864
Db	1873	GGTATTAGCATGATGTTGGAAATTAAGAAAGTCAATTTACAGAGAGAGGTTTAAAGTTGGTA	1932
Qy	1865	TTGGTCAATCTCTGTAAGTGAAGTGAAGAAACTGAAGAACTGAACAAATCGAAGT---TCCAAAT	1921
Db	1933	TTGTCAATCCAAAGAGAGAGGTGCGTGAAGAAATTAACCAAGATCCAAATTCATCGGTGAT	1992
Qy	1922	CATTTAGGGAAGAAATGATCTATCTGACTGTGTTGAAGAGCCGTTGGAGCATGCAACTTC	1981
Db	1993	CATTTGGGCAAGAGTGGATGTTTTCACGGTAGGAGAGAGCAGTGGAGGCTTGTAGCTAC	2052
Qy	1982	AATCTACGTGCAAGCAAAACGAAACC 2007	
Db	2053	ATGCTTCAACAGCTTTTAAACCGAACC 2078	

RESULT 4
ABZ13697
ID ABZ13697 standard; DNA; 1977 BP.
XX ABZ13697;
AC ABZ13697;
DT 21-JAN-2003 (first entry)
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1502.
DE Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN WO200216555-A2.
XX 28-FEB-2002.
PD 24-AUG-2001; 2001WO-US026685.
XX 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0266467P.

Db 1737 CGGTATTAGCATGATGTGGAAATTAAGAAAGTCATTTGACAGGAGCGGTAAAGTTGGT 1796
Qy 1864 TTTGGTCAATCTCTGAAGTGAAGTGAAGAACTGAAGAAATCGAAGT---TCCAAAA 1920
Db 1797 ATTGTCAATCCAAAAAGAGAGGTCTGTGAAGAAATTAACAGATCCAAATTCATCGGTGA 1856
Qy 1921 TCATTAGGAGAAATGGATCTATCTGACTGTGTAAGAGCGCGTTGGAGCATGCAACTT 1980
Db 1857 TCATTTCGGCAAGAGTGGATGTTCTTAACGGTAGGAGAGCAGTGGAGGCTTTGTAGCTA 1916
Qy 1981 CAATCTACGTGCAAGCAAAAGCAACCC 2007
Db 1917 CATGCTTCACACGTTTAAACCGAACC 1943

RESULT 5
ID ADA68514
AD ADA68514 standard; DNA; 1977 BP.
XX
AC ADA68514;
DT 20-NOV-2003 (first entry)
XX
DE Arabidopsis thaliana gene, SEQ ID 589.
XX
KW Plant; bacterial infection; fungal infection; viral infection; ds.
XX
OS Arabidopsis thaliana.
XX
PN W02003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 589; 899pp; English.
XX

CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX

SQ Sequence 1977 BP; 513 A; 458 C; 471 G; 535 T; 0 U; 0 Other;
Query Match 41.68; Score 1018.2; DB 8; Length 1977;
Best Local Similarity 72.3; Pred. No. 1e-240;
Matches 1364; Conservative 0; Mismatches 518; Indels 5; Gaps 3;
Qy 125 CACCAAGTCGAGGTTCACCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG 184
Db 58 CACAGGTGGAGGCTCCACAACCTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAG 117

Qy 185 GAGACTTTCTTCCCTGATGACCCCTTTGAGGAGTTCAAGAACAGCCAGCTTCCAAAGAG 244
Db 118 GAAACTTCTGTGTTCCAGACGACCCCTTTTAGACAAATTTAAGAACAAATATGATCAAGAAA 177
Qy 245 TTCATGCTTGGCCCTTCAGATTCTTCTTCCCATTTTCGAAATGGGCTCCCAATACACCTTT 304
Db 178 TTTGTGTAGGCTCAAAATACTTCTTCCGATTTTGAATGGGACCAACGCTACATCTC 237
Qy 305 CAGTTCTTGAAGCTCACTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAG 364
Db 238 AAGTTCTTCAAAATCAGATCTCATCGCGGAATCACCATCGTAGTCTCGGCATCCCTCAG 297
Qy 365 GGCATCAGTTATGCCAAGCTCGCAACCTCCCTCCGAATCTTGGACTATATTCAGAGTTT 424
Db 298 GGCATCAGTTATGCCAAGCTTGTAACTTGCCTCCCAATCTTGGCCCTTTATTCAGATTT 357
Qy 425 ATACCACCAATTGATTTATCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGAGCTGTG 484
Db 358 GTACCGCCATTTGGTATACGGGTGCTAGGAGTTCAAGGACTTAGCGGTGGGAACGGTT 417
Qy 485 GCGTTTGGATCGCTTCTGATGGGTTTCGATGTTGAGTAATGCCGTTGATCCCAATGAAGAC 544
Db 418 GCGTTTGGCTCTCTGTTGACAGGTGCGATGCTGAGCAAGAAAGTTGATGCTGAGAAAGAT 477
Qy 545 CCAAAGCTTTACCTCCACCTGGCTTTCAGACTACATTAATTTGCTGGTGT-TTTCAGGC 603
Db 478 CTTAAGCTTTACCTTCCACCTTGTCTTTCACCGCCACTTTTTTCGCCGCGGTTCTCGAAGCC 537
Qy 604 TGCCTTGGGTCTGTTTATAGTTGGGTTGATCGTGGATTTTCTGTCACATCAACCAATAAT 663
Db 538 TCTCTTGGAAATTTTCAAGGTTAGGGTTTCAATAGTGGACTTTCTATCGATGCAACGATAGT 597
Qy 664 AGGTTTCATGGGAGGAGCAGCCACCGTGGTGTCTGACGAACTAAATTCGATTCGTTGG 723
Db 598 AGGATTCATGGGAGGAGCAGCGAGTGTGAGTCTGCAACAGCTTAAAGGATTTTTCGG 657
Qy 724 CTTGAGCATTTTCAACCATGGAGCTGATATCATATCAGTATGCGCTGTTTTCACCCA 783
Db 658 ACTTAAACATTTTACAGACTCTACCGATGTTATCTCTGTCATGCGTTCGTTTTCTCCCA 717
Qy 784 AACTCATAGTGGAGGTGGAAAGTCTGTGTTAGGATGTTCTTCAATTTCTTCTCTCCT 843
Db 718 AACTCAGAGTGGAGATGGAAAGTGGCGTCTTGGATGGTGTCTTATCTTCTTCTTCT 777
Qy 844 TAGCACAAGATCTTACGCAAAAAACGCAAGGTTTTTTGGGTGTGACAAATGCGCC 903
Db 778 CTCACACAGATTTTGAAGCATCAAGAAACCAAAATTTCTTTGGGTGGCGGATGGCTCC 837
Qy 904 ATTGAGCTCGGTATATTTGGGAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGACGG 963
Db 838 TTTGACCTCAGTGATTTCTTGGAAAGTCTCTTGGTTTACTTCTCCTCAGCCTGAGAGATGG 897
Qy 964 TGTGAAGTATAGGAGAACTGAAGAAAGGTTTGAATCCACCATCACTCACAATCTGGT 1023
Db 898 TGTTCAGTATAGGAGACCTGAAGAAAGGTTGAATCCACTCTCGGTTCTGATCTCAT 957
Qy 1024 ATTTGTGTGCGCTTACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATCATATC 1083
Db 958 CTTTACTTCCCTTACATGTCACAGCTGTCAAACTGGCCTCATCACTGGCATCATTCG 1017
Qy 1084 ACTTGGGAAGGATAGAGTAGGAGAGAGCTTTTGAATGTATAAAATTTACAATATGA 1143
Db 1018 TCTCGC-GAAGGAGTAGCTGGGGAGGAGTTTTTGGGATGTTTCAAGAACTTCAACATAGA 1076
Qy 1144 TGCAACAAAGAGATGATAGCTATTGGGACCAAGACGTTAGTTGGTTCTTTCACCTCTTG 1203
Db 1077 CGGGAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTTGGTTCTTTCACATCTTG 1136
Qy 1204 CTACCTCACAACAGGACCACTTTTCGCGTTTCGGGTGTGAACACTATAACGCTGGATGCAAGAC 1263
Db 1137 TTACCTCACAACCGGACCACTTTTCAAGGTCGGCAGTGAACACTACAACGCGTTTGAAGAC 1196
Qy 1264 AGCAGCTTCAACATTTATTAATGTCACCTTCGAGTAATGTTGACATTTGTTATCTCTGACACC 1323

Db 1197 CGCAATGTCACATAGTAGGATGCGGTTATGTTTCACACTCTCTCTCTCACACC 1256
Qy 1324 CTTGTTCCATTACACTCCCTGGTGGCTATCAGCTATTATCGTATCGCAATGCTGG 1383
Db 1257 GCTTTTCTACTACACACCACTCGTGGCTCTCTGCCATCATCATATCGCAATGCTCG 1316
Qy 1384 ACTCATAGATTATGAAGCAGCCATCCATCTATTAAAGTTGACAAATTTGACTTTGTGT 1443
Db 1317 ACTCATGACTATCAAGTGCATCCATCTCTGGAAGTTGACAGTTGCACTTCTCGT 1376
Qy 1444 GTGCATGAGTGCATACATTGGCGTGGTCTTTGGCAGTGTGAAATGGCTTAGTCATAGC 1503
Db 1377 CTGCATGAGCGCTACGTTGGGTCGTATTCCGCGAGTGTAGAGATTGGACTCGTCTAGC 1436
Qy 1504 TATTGTAATCTGTACTCTGGGTACTCTTATTATGCAAGGCCAAGACATTCGTTTT 1563
Db 1437 GTTGGCGATATCTATACGAGGTTGTGCTGTTGTGTCAGGCCAAAACTCGGTGAA 1496
Qy 1564 GGCAACATTCCAAATTCCTGTATATACCGAAATGTTGACACTATCAAAATGCAAAACA 1623
Db 1497 GGAACATACCAACAGCATGATCTATAGAACACTGACGATCCCATCATCAAGAAC 1556
Qy 1624 TGTCTCTGGAATGCTAATCTAGAGATTGATGACCAATTTACTTTGCCAATGCCAGTA 1683
Db 1557 CGTTCCTGCTATTCTCATCTTGAGATGTATGCTCCCATCTACTTTGCTAAATGCCAGTTA 1616
Qy 1684 TTAAAGAAAGGATCACAGGTGGATGATGATGAAGAGAGAAAGATTAAAGCTACAGG 1743
Db 1617 CTTGCGTGAAGAAATCATAGGTGGATTGATGAAGAGAGAGAGATTAAACAATCAGG 1676
Qy 1744 GGAGACTAGTTTGCAGTATCTTATAATTGATATGAGTGTGTTGGAAACATTCATACAAG 1803
Db 1677 AGAGAGCAGTTTCAATATATATATCTCGATATGTCAGCTTGTGTTGTTATATCGACAAAG 1736
Qy 1804 TGAATAAGTATGCTTGAAGAGGTGAAGAGATTATACAGAGAGAGAGAGCTACAGCTTGT 1863
Db 1737 CGGTATTAGCATGATGTTGGAATTAAGAAAGTCATTGACAGGAGCGTTAAAGTTGTT 1796
Qy 1864 TTTGCTCAATCTGTAAGTGAAGTATGAGAACTGAGAACTGAACAAATCGAAGT---TCCAAA 1920
Db 1797 ATTGTCAAATCCAAAAGGAGAGGTGCTGAAGAAATTAACAGATCCAAATTCATCGGTGA 1856
Qy 1921 TCATTAGGAAGAAATGGATCTATCTGACTGTTGAAGAGCGGTTGGAGCATGCAACTT 1980
Db 1857 TCATTGGGCAAGAGTGGATGTTCTTAACGTTAGGAGCAGTGGAGGCTTGTAGCTA 1916
Qy 1981 CAATCTAGTGCAGCAAAACGAACCC 2007
Db 1917 CATGCTTCACACGTTTAAACCGAACC 1943

RESULT 6

ADAT0002
ID ADA70002 standard; DNA; 1974 BP.
XX
AC ADA70002;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 3325.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX

PR 22-JUN-2001; 2001WO-IB001105.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
DR
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX Claim 6; SEQ ID NO 3325; 899pp; English.
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 1974 BP; 335 A; 602 C; 657 G; 380 T; 0 U; 0 Other;
Query Match 34.5%; Score 846; DB 8; Length 1974;
Best Local Similarity 66.4%; Pred. No. 3.3e-198;
Matches 1231; Conservative 0; Mismatches 620; Indels 3; Gaps 1;
Qy 131 GTCGAGGTTCCACGCCACAGCCGCTTTTCAAGTCTCTTAAAGTACTCTTTGAAGGAGCT 190
Db 55 GTGCCGATGCCCGCGCGAAGCCGTTCTCTGGAGACGCTGGGGGGGAACAATGAAGGAGACA 114
Qy 191 TTTCTTCCCTGATGACCCCTTTGAGGACAGTTTCAAGAACAGCAGCTTCCAAAGATTTC--- 247
Db 115 TTCTCTCCGAGACACCCGTTTCAAGTGTGTGCGCGGAGCGGGTGGCGCGCGCG 174
Qy 248 ATGCTTGGCCTTTCAGTCTTCTTCCCAATTTTCGAATGGGCTCCCAAAATACACCTTTTACG 307
Db 175 GCGCGCGGCTCCGCTACGTTTCCGTTTCATGAGTGGGCGCGCTGCTACACCTCCGCG 234
Qy 308 TTTCTTGAAGTCTGCTCATAGCTGCATCACCATCGCTAGCTTGGCCATTCCTCAGGGC 367
Db 235 ACCCTCAAGTCCGACCTCATCGCGGATCACCATTGCGAGCTCGCCATCCCCCAGGGC 294
Qy 368 ATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATTTCTTGGACTATATTCGAGCTTTATA 427
Db 295 ATCAGTACGCCCAAGCTCGCCAACTCCCTCCGCTCTCTATTCGAGCTTCGNG 354
Qy 428 CCACCAATGATTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGGACTGTGGCG 487
Db 355 CCGCGCTGCTGTACGCGATGATGGGAGCTCGAGGACCTCGCGGTGGGACCGTGGCG 414
Qy 488 GTTGGATCGCTTCTGATGGGTTGATGTTGAGTAATGCCGTTGATCCCAATCAAGACCCA 547
Db 415 GTGGCGTCTGCTGATCGGGTCTGATGTCGCGAGGAGGTGTCGGCGCGGAGGACCCG 474
Qy 548 AAGCTTTTACCTCCACTGGCTTTTACAGCTACATTTATTTGCTGGTGTGTTTTTCAGGCTGCC 607
Db 475 GCGCTGTACTGACGCTCGCGCTACCGCCACCTTTCTTCGCCGCGCTGTTTCAGGCGCTG 534
Qy 608 TTGGGTCTGTTTAGGTTGGGGTTGATCGTGGATTTTCTGTACATGCAACCAATATAGGG 667
Db 535 CTGGGGGTCTTCAGGCTGGGGTTTCATCGTGGACTTCTCTGTGCGACGCCACCATCGTCGG 594
Qy 668 TTCATGGGAGGAGAGCCACGCTGTGTCTGCGAGCAACTAAATTCGATTTCTTGGCCTT 727
Db 595 TTCATGGGCGCGCGCCACCGCTGTGTGCTCTGACAGCTCAAGGGCATGTTGCGCCTC 654

```
Qy 728 GAGCATTTCCACCATGAGCTGATATCATATCAGTATGGCTCTGTGTTTTCACCCAACT 787
Db 655 GACCACCTTCCACCGCCACCGACCTCGTCTCCGTCATGAGCTCCGCTCTTCTCCCAACC 714
Qy 788 CATGAGTGGAGCTGGGAAAGTGTGTAGTAGTGTCTTCAATTTCTTCTCTCTTAGC 847
Db 715 CACCTCTGGCGATGGGAGAGCGTGTCTATGGCTGGGCTTCTTCTTCTCTCTCATC 774
Qy 848 ACAAGATATTTACAGAAAAACGACCAAGTTTTTTTGGGTGTCAGCAATGGGCCATTG 907
Db 775 ACCCGCTTCTTACAGAGAGAGGAGGCAAGGTTCTTCTGGGTATCTGCAGCTGGGCCATTG 834
Qy 908 ACCTCGCTTATTTGGGAGGCTCTCTGTTTATTTTCACTCAGCGCGAGAGACGAGTGT 967
Db 835 CGCTCTGTCTATCTCGGGAGCTCTCTGGTGTACTCACACATGTGAAACCATGGCATT 894
Qy 968 GAAGTGTAGGAGAACTGAAGAAGGTTTGAATCCACCATCACCTCACAAATCTGTATTT 1027
Db 895 CAAGTGATTGGTTACCTGAAGAAGGCTGNAACCCACCGTCTGCCACAAGCCTGAACCTC 954
Qy 1028 GTGTGCGCTTACATGACTACAGCTGTCAAACTGGCAATTTGTGTGGATCATATCACTT 1087
Db 955 TCGTTCGCGTACATGATGTGCGCTTGAAGACCGGATCATCACTGGCGTCAATTCGCCCTC 1014
Qy 1088 GCGGAAGGATAGCAGTAGGAGAGCTTTGCAATGTATAAATTTACAATATTGATGGC 1147
Db 1015 GCTGAGGGATGCGCGTAGGAAGAGCTTTTCAATTTTCAAGAAATTACCAATCGACGGA 1074
Qy 1148 AACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTGTGTTCTTTTCACTCTTGCTAC 1207
Db 1075 AACAAAGAGATGATCGGCTTTGGAAACGATGAACATTTGTGGATCACTCACTCTTGCTAC 1134
Qy 1208 CTCACACAGGACCAATTTTCGCGTTCGGCTGTGAACATATAACCGTGGATGCAAGACAGCA 1267
Db 1135 CTCACACCGCTCGGTTCTCGAGTTCGGCGTCAACTACAAACGCGCGTCAAGACGCGG 1194
Qy 1268 GCTTCCAACTATTAATGTCACTTGCAGATGATGTGATGATTTATTCCTGACACCTTG 1327
Db 1195 ATGTCGAACGTGATCATGTTCGGTGGCGGTGATGATCACTGCTGTGTTCTTGACGCGCTG 1254
Qy 1328 TTCCATTTACACTCCCTCGGTGCTCTATCAGCTATTATTCGATGTCGAATGCTTGGACTC 1387
Db 1255 TTCCACTACACCGCTGCTGTGCTGTGCGGATCATCATGTGCGGATGCTGGGCTG 1314
Qy 1388 ATAGATTATGAAGCAGCATTCATCTATTTAAGTTTGACAAATTTGATTTGTGTGTC 1447
Db 1315 ATCGACTACCGCGCGCTCCACCTGTGGCAGGTGACAAAGTGGACTTCTGCTCTGC 1374
Qy 1448 ATGATGATACATATGGGTGCTTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTATT 1507
Db 1375 CTGGCGCTTACCTCGGCGTGTCTTTCGCGAGCGTTCAGATCGGCTTCGTCGCGCTC 1434
Qy 1508 GTAATATCTGACTTCGGGTACTTCTATTTATTCGAAGGCCAAGGACATTCGTTTGGGC 1567
Db 1435 GGGATCTCCATCTCCGCGTGTCTGTTCGTGGCGCGCGGAGGACGAGTGTCTCGG 1494
Qy 1568 AACATTCCAAATTTGTGTATATCCGAAATTTGTGAGCATATTCAAATTCGAAACATGTT 1627
Db 1495 AACATCCCCAACAGATGATCTACCGCGGATGACAGTACACCGCGCGCAGAGGCTC 1554
Qy 1628 CTTGAATGTATTTCTAGAGATTGATGACCAATTTACTTTGCAATGTCAGTATTTA 1687
Db 1555 CCGCGGTGCTCGTGTCTTCGGCTGCACTTCGCCCATCTACTTACCAACGCCGAGTCACTG 1614
Qy 1688 AGAGAAAGGATCACAGGTGATGATGAAGAAGAAAGAAATTAAGCTACAGGGGAG 1747
Db 1615 CGTGAGAGATTCGCGGTGATGATCGACGAGGAGGACGATGTCAGAGGAGAGGGCGAG 1674
Qy 1748 ACTAGTTTGCAGTATGTTTATAATTTGATAGTGTCTGTGGAACATTTGATACAGTGA 1807
Db 1675 ATGGGCATTCAGTACGTCGCTCTCGACATGGTTCGGTTCGAGCATCGACACGCGGG 1734
Qy 1808 ATAAGTATGTTGAAGGTGAAGAAGATTACAGAGAGAGAGACTACAGCTTGTGTTTG 1867
```

```
Db 1735 ACAGCATGCTGATGAATCAAGAGACCTCGACAGGAGGCGCTTCAGATTGTGTTG 1794
Qy 1868 GTCAATCTCTTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTTCCAAATCATTTA 1927
Db 1795 GCGAACCCGGGAGCGAGATCATGAAGAAGTTGGACAGCTCCAAAGTGTCTTGAGCGCATC 1854
Qy 1928 GGAAGAANAATGATCTATCTGACTGTGTTGAAGAGCGCGTTGGAGCATGCAACTTC 1981
Db 1855 GGCCATGAGTGGATCTTCCACCGTGGCGAGCGGTGGCGGAGTGCAGACTTC 1908

RESULT 7
AAZ50482
ID AAZ50482 standard; cDNA; 1981 BP.
XX
AC AAZ50482;
XX
DT 23-MAY-2000 (first entry)
XX
XX Corn sulphate permease-2 cDNA clone.
XX
XX Sulphate Permease; sulphate assimilation protein; corn; contig; probe;
XX mapping; marker; plant breeding; chimeric gene; transgenic plant;
XX antibody; screen; ss.
XX
XX Zea mays.
XX
XX Key Location/Qualifiers
XX CDS 2..1741
XX
XX FT /tag= a
XX FT /product= "Corn sulphate permease-2"
XX FT /note= "Derived from contig composed
XX FT crln.pk0015.a2, p0006.cbyvs25rb, p0072.comhc25r,
XX FT p0091.cmard29r and p0092.chwat43r"
XX
XX PN WO200004154-A2.
XX
XX PD 27-JAN-2000.
XX
XX PF 13-JUL-1999; 99WO-US015810.
XX
XX PR 14-JUL-1998; 98US-0092833P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX PA Allen SM, Falco SC, Thorpe CJ;
XX
XX PI WPI; 2000-195025/17.
XX
XX DR P-PSDB; AAY44936.
XX
XX Nucleic acid fragments encoding sulfate assimilation proteins in plants
XX PT and seeds useful as probes for isolating cDNAs and genes encoding
XX PT homologous proteins, in producing transgenic plants.
XX
XX PS Claim 3; Page 41; 79pp; English.
XX
XX The present sequence is a cDNA clone encoding corn sulphate permease, a
XX sulphate assimilation protein. This sequence is from a contig composed of
XX clones crln.pk0015.a2, p0006.cbyvs25rb, p0072.comhc25r, p0091.cmard29r
XX and p0092.chwat43r from crln, p0006, p0072, p0091, and p0092 libraries,
XX respectively. This sequence is used as a probe to isolate other plant
XX sulphate assimilation proteins, for genetic and physical mapping of
XX related genes and as markers of traits linked to the gene. This is useful
XX for plant breeding and to construct chimeric genes, used to create
XX transgenic plants with altered levels of sulphate permease. The sulphate
XX permease peptides are useful for producing antibodies, that are used to
XX screen and isolate cDNA clones
XX
XX SQ Sequence 1981 BP; 388 A; 621 C; 612 G; 360 T; 0 U; 0 Other;

Query Match 31.2%; Score 765.2; DB 3; Length 1981;
Best Local Similarity 65.8%; Pred. No. 2.9e-178;
```

Matches 1112; Conservative 0; Mismatches 578; Indels 0; Gaps 0;		
QY	313	GAAGAGTACCTCATAGCTGGCATCACATCGCTAGCTTGGCGCATTCCTCAGGGCATCAG 372
Db		
QY	7	GGAGTCGACCTGATCGCGGCATCACATCGCAGCCTCGCCATCCCGCAGGGCATCAG 66
Db		
QY	373	TTATGCCAAGCTGCCAACCTCCCTCCAAATCTTTGGACTATATTCGAGCTTTATACACC 432
Db		
QY	67	CTAGCCCAAGCTGCCAACTCGCGCCGCGTCTCGACTCTACTCGAGCTTCGTGCGCC 126
Db		
QY	433	ATTGATTTATGCGATGATGAGTGTAGCTCGAGGGATTTGGCAGTGGGACTGTGCGGTTGG 492
Db		
QY	127	GCTGGTGTACCGCTGATGGGGAGCTCCAAAGACCTGGCGTGGGACCGTGGCGTGGC 186
Db		
QY	493	ATCGCTTCTGATGGTTCGATGTTGAGTATGCGTGTATCCCAATGAAAGACCCAAAGCT 552
Db		
QY	187	GTGCTGCTCATCAGCTTCCATGCTCGGCACGAGGTGTGCCGACGGAGAACCCCGTGCT 246
Db		
QY	553	TTACCTCCACCTGGCTTTACAGCTACATTTATTTGCTGGTGTGTTTTTCAGGCTGCCTTGGG 612
Db		
QY	247	CTACCTGCACCTCGCCTTACCGCCACCTTCTTCGCGCGGTCTTCCAGGCTCGCTCGG 306
Db		
QY	613	TCGTTTAGTGGGTTGATCGTGAATTTCTGTCAATGCAATGCAATCAATAGGGTTCA 672
Db		
QY	307	CCTCCTCAGGTTGGGCTTCACTCGTGACCTGCTGTGCGACGACGATCGTGGGGTTCA 366
Db		
QY	673	GGGAGGAGCACCACGCTGGTGTCTGCAGCAACTAAATCATTTCTTGGCCCTTGAGCA 732
Db		
QY	367	GGCGCGCGGCGACGGTGTGTGCTGCAGAGCTGAAGGGCATGCTGGGCGCTCGTCCA 426
Db		
QY	733	TTTCACCCATGAGAGCTGATATCATATCAGTGATGCGCTCTGTTTTTCAACCCAACTCATGA 792
Db		
QY	427	CTTCACCACTCCACCGACGCTGCTCCCGTCATGGAATCGTCTTCAGCCAGACACCA 486
Db		
QY	793	GTGAGGTGGGAAAGTGTGTTAGGATGTGTCTTCAATTTTCTTCTCTTTAGCAAG 852
Db		
QY	487	GTGGCGTGGGAGAGGTCTGCTGCGCTGCGGCTTCTCTTCTCTCTCTGCTCACCG 546
Db		
QY	853	ATACTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGCGCGCATTTGACGTC 912
Db		
QY	547	CTTCATCAGCAAGGGTCCCAAGCTGTTCTGGATCTCGCGGCGCGCGTTGACGTC 606
Db		
QY	913	CGTTATATTGGGAAGTCTCTTGGTTTATTTCACTACCGCGAGAGCAGGTTGGAAGT 972
Db		
QY	607	CGTCGTGCTCGGAGCGTCTCTGTGTACCTCACGACGCTGAAACCAACGGCATCGAAGT 666
Db		
QY	973	GATAGGAACTGAAGAGGGTTGAATCCACCATCACTCAAAATCTGGTATTTGTGTC 1032
Db		
QY	667	GATCGGTTACTGAAGAAAGGCTGAATCCACCGCTCGGTGACAAAGCTGCAATTTCTCACC 726
Db		
QY	1033	GCCTTACATGACTACAGCTGTCAAACTGSCATTTGCTGGCATCATATCACTTCGCGGA 1092
Db		
QY	727	GCCTTACATGATGCTCGCGCTCAAGACTGGGATCATCACCGGCTCATTTGCCCTCGCCGA 786
Db		
QY	1093	AGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAATTAACAATTTGATGGCAACAA 1152
Db		
QY	787	AGGAATCGCGTGGGAGGAGCTTCGCCATGTTCAAGAACTACCAATGACGACACAA 846
Db		
QY	1153	AGAGATGATAGCTATTGCGGACCAATGAACGATGTTGTTCTTTCACCTTTGCTACTCAC 1212
Db		
QY	847	GGAGATGATCGCATCGGACGATGAACGTCCTGGGCTCGCTCACGCTCGTGTACCTGAC 906
Db		
QY	1213	AACAGACCACTTTTCGGCTTCGGCTGTGAACATAACGCTGGATGCAAGACAGCAGCTTC 1272
Db		
QY	907	CACGGGGCCCTTCTCGCGCTCCCGGTGAACACTAACCGCGGTGAGGACGGCCCATGTC 966
Db		
QY	1273	CAACATTATAATGTCACTTGCAGTAATGTTGACATTTGTTATTCCTGACACCCCTTGTTC 1332
Db		
QY	967	GAACGTGGTTCATGCTGCTGCGGTGATGGTTCACGCTGCTTCTTCTGACCGCGCTTCCA 1026
Db		
QY	1333	TTACACTCCCTGGTGTGCTATCAGCTATTATCGTATCTGCAATGCTTGGACTCATAGA 1392
Db		
QY	1027	CTACACCGCGCTGGTGGTGTGCTGTCGGCGATCATCGTCTCGCGATGCTGGGCTGTGTCGA 1086
Db		

RESULT 8
AAZ50488
ID AAZ50488 standard; cdNA; 2067 BP.
XX
AC AAZ50488;
XX
DT 23-MAY-2000 (first entry)
XX
DE Soybean sulphate permease-1 cdNA clone.
XX
DE DE
XX
KW Sulphate Permease; sulphate assimilation protein; soybean; probe;
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
XX antibody; screen; ss.
XX OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 2..1867
FT /product= a
FT /note= "Soybean sulphate permease-1"
FT /note= "Derived from clone ses2w.pk0031.b3"
XX
PN WO200004154-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015810.

XX	14-JUL-1998; 98US-0092833P.	QY	743	GGAGCTGATATCATATCAGTATGCGCTCTGTTTTCACCCAAACTCATGAGTGGAGGTGG	802
XX	(DUPO) DU PONT DE NEMOURS & CO E I.	DB	608	AAGATGCAAAATAATCCAGTAAACGATCTCTGTTTTCAGCAAGACACAGAGTGGTCATGG	667
XX	Allien SM, Falco SC, Thorpe CJ;	QY	803	GAAAGTGTGTGTAGGATGTGCTTCAATTTCTCTCTCTCTAGCACAAGATATCTTCAGC	862
XX	WPI; 2000-195025/17.	DB	668	CAAAACCATCTTTTGGGATTCGGCTTCTGCTCTCTCTGCTGACCAAGGACATATAGT	727
XX	P-PSDB; AAY44942.	QY	863	AAAAAAGCAACCAAGTGTTCAGCAATGGCCATTCAGCGTTCGGTATATATG	922
XX	Nucleic acid fragments encoding sulfate assimilation proteins in plants	DB	728	TTGAGGAAACCAAAATATTTCTGGGTTTTCAGCAGCTGCCCAATGACATCAGTATATCTG	787
XX	and seeds useful as probes for isolating cDNAs and genes encoding	QY	923	GGAGTCTCTTGGTATTTTCACTCAGCGGAGAGCAGGCTTGAAGTATAGGAGAA	982
XX	homologous proteins, in producing transgenic plants.	DB	788	TCAACCATTTTGTCTTCTCTGAGAAATAAGACTCATCAAAATTTCAAGTATTTGGGCAC	847
XX	Claim 3; Page 49-50; 79pp; English.	QY	983	CTGAAGAAGGTTTGAATCCACCATCCTCACAATCTGCTATTTGTGTCCTTACATG	1042
XX	The present sequence is a cDNA clone encoding soybean sulphate permease,	DB	848	TTACCAAGGGAGTTAATCCACCATCAGCAACATGTTATCTTCAATGGTCTCTTACTTG	907
XX	a sulphate assimilation protein. This sequence is obtained from ses2w	QY	1043	ACTACAGCTGTCAAAAACCTGGCAATGCTGGTGGCATCATATCACTTCGGGAAGAAATAGCA	1102
XX	library, clone ses2w.pk0031.B3, derived from soybean embryogenic	DB	908	GGTCTTGTCTATCAAAAACCTGGCATCATCAGAGGATCTTATCTCTCACTGAAGGAATGCA	967
XX	suspension 2 weeks after subculture. This sequence is used as a probe to	QY	1103	GTAGGAAGAGCTTTCGAATGTATAAAATTTACAATATTGATGGCAACAAAGAGATGATA	1162
XX	isolate other plant sulphate assimilation proteins, for genetic and	DB	968	GTAGGGAGAAATTTGCTTCTTCACTTAAGAACTACAGGTGGATGGAACAAAGAAATGATG	1027
XX	physical mapping of related genes and as markers of traits linked to the	QY	1163	GCTATTGGGACCATGAACGTAGTGTGTTCTTTCACCTCTTGTCTACCTCACAACAGGACCA	1222
XX	gene. This is useful for plant breeding and to construct chimeric genes,	DB	1028	GCAATTTGGTCTAATGAACATAGCTGGCTGTGTTCTTCTATGTTATGTTACAGGGATCC	1087
XX	used to create transgenic plants with altered levels of sulphate	QY	1223	TTTTCCGCTTGGCTGTGAACCTATAACCGCTGGATGCAAGACAGCAGCTTCCAAATATTA	1282
XX	permease. The sulphate permease peptides are useful for producing	DB	1088	TTTTCTCGATCGCTGTAACTATAATGCTGGAGCAGACACAGTTTCAAATATAATC	1147
XX	antibodies, that are used to screen and isolate cDNA clones	QY	1283	ATGTCATCTGCAGTAATGTTGACATTTGTTATTCCTGACACCTTGTTCACATACATCCC	1342
XX	Sequence 2067 BP; 583 A; 450 C; 393 G; 641 T; 0 U; 0 Other;	DB	1148	ATGGCTGCACTGCTTCTAGTGACACTTCTGTTTCTCATGCTCTTCTTACTATACACCA	1207
XX	Query Match 29.48; Score 720.6; DB 3; Length 2067;	QY	1343	CTGGTGGTGTATCAGCTATATTCGATCTGCAATCTTGGACTCATAGATTTATGAAGCA	1402
XX	Best Local Similarity 62.24; Pred. No. 3e-167;	DB	1208	AATGTTGCTTAGCGGCCATTTATCATCATCTGCTGATGTTGTTATAGATTTATCAATCT	1267
XX	Matches 1134; Conservative 0; Mismatches 689; Indels 0; Gaps 0;	QY	1403	GCATCCATCTATTTAAGGTGACAAATTTGACTTTTGTGTGTGTCATAGTGCATACAT	1462
QY	143 CCGCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACTTTCTTCCCTGAT	DB	1268	GCATATAAATTTGGAAGGTGACAAATTTGATTTCTTGGCCTGTTGTGCTCCTTTTTT	1327
DB	8 CCACACGACCACTCCACAACTTAGGCACAGAGCTCCGAAATCTTCTCCCGAGAT	QY	1463	GGCGTGGTCTTGGCAGTGTGAAATTTGGCTTAGTCTAGTCTATGTAATATCTGTACTTT	1522
QY	203 GACCTTTGAGGAGTTCAGAAACAGCGAGCTTCCAAAGATTTGCTTGGCTTTTCAG	DB	1328	GGGGTCTGTTCATTTTCACTGCTTGTAGGTCTTGGTATAGCGGTATTCATATCAGTCCCTC	1387
DB	68 GACCTCTCCACCGTTTTTCAAGAACCAAACTCGCTTTTAAAGATTTCTCTCGCACTTCAG	QY	1523	CGGGTACTTCTATTTATTCAGGCGCAAGGACATTCGTTTGGGCAACATTTCCAAATTC	1582
QY	263 TTCTTTCTCCCAATTTTGAATGGCTCCCAATATACACTTTTCAGTCTTTGAAAGCTGAC	DB	1388	AGATCTGCTTCATGCTACTCGACCAACACTTTGCTTGGTGTGGGGAATATACCAAGNACA	1447
DB	128 TATCTTTTCCCAATTTTTCGACTGGGCCCCAACTACAAATCTTACCTTCTCCGCTCTGAC	QY	1583	GTGATATCCGAAATTTGAGCACTATCAAAATTCAAAATTCGCAATCTCTTCCGAAATTC	1642
QY	323 CTATAGCTGCATCACCATCGCTAGCTTGGCCATTCCTCAGGGCATCAGTTATGCCAAG	DB	1448	CAAAATTTCCACAACATAAACCAATAACAAAAGCTTTAAGAGTCTCTTCAATTTCTCAT	1507
DB	188 CTATCTCTGGCTCACCATTGCCAGCTCGCCATTCCTCAGGAATCAGTTATGCCAAG	QY	1643	CTAGAGATTTGATGACCAATTTTACTTTGCCAATGCCAGCTATTTAAGAGNAGAGATCACA	1702
QY	383 CTGCCAACCTCCCTCCAAATCTTTGGACTATATTCGAGCTTTTATACCAATGATTTAT	DB	1508	TTGGCTGTTGAGTCTCCAATCTATTTTGTCTAACTCACTTATCTTCAAGAAAGGATACTG	1567
DB	248 CTTGCCAACTTGCACCTATTTCTGGATTTATTCGAGTTTGTTCGCCCATTTGATATAC	QY	1703	AGGTGATTTGATGAAGAAGAAAGAAATTTAAAGCTTACAGGGGAGACTAGTTTGCAGTAT	1762
QY	443 GCGATGATGGTACTCGAGGATTTGGCAGTGGGACTGTGGGTTGGATCGCTTCTG	DB	1568	AGATGGGTTTCGAGNAGAGGAGAGCATATAAAGCTTAATAATGGAGCTCCATTTGAAGTGC	1627
DB	308 TCGCTGTCTGGAATTTTACACATCTTGGTGTGGACCTGTGTTTCAATTCGCTCTTGGTC	QY	1763	GTATTAATTTGATAGTGTCTGTTGAAACATTTGATACAAGTGGAAATAGATGCTTGAA	1822
QY	503 ATGGGTTTCGATGTTGAGTAAATGCGGTTGATCCCAATGAAGACCCAAAGCTTTTACCTCCAC	DB	1628	ATAATTTTAGACATGACAGCTGTCCAGCCACAGACACAGTGGCTTGACACTTTATGT	1687
DB	368 ATGGGATCAATGTTAAGTGATATAAATTTCTTACACTCAAGNACCTATTTCTATCTGGGA	QY	1823	GAGGTGAAGAAGATTCAGAGAGAAGAGACTACAGCTTGTGTTTGGTCAATCTCTGTNAGT	1882
QY	563 CTGGCTTTCAGACTACATTTATTTGCTGGTGTGTTTTTTCAGGCTGCTTTGGTCTGTTTGG	DB			
DB	428 TTGGCTTTCACCGCACTTTCTTGTCTGGTGTATTTCCAAGCTTCTCTGGGTATATTAAGG	QY			
QY	623 TTGGGTTTGTATGCTGATTTCTGTCATCGACCAACATAATAGGTTTCATGGGAGGACA	DB			
DB	488 CTAGGCTTCGTAATTTGATTTTCTGTCGAAGCAACGCTGTTGTAATTCACAGGCGGTGCT	QY			
QY	683 GCCACGGTGGTGTCTGTCGACCAACTAAATTCGAATTTCTTGGCTTGTAGCACTTTCAACCAT	DB			
DB	548 GCCATTAATTTGTCTACCTGCAGCAGCTGAAGAGTTTACTTTGGAATAGTGCACTTTACGAC	QY			

Db	1688	GAAGCTTAGAAGATGCTGGAGAGAGATCACTTGCAGTTTGTGCTGGCAAAATCCTGTTGGA	1747	PR	08-JUN-1999;	99US-0138094P.
				PR	10-JUN-1999;	99US-0138540P.
Qy	1883	GAAGTGATGAAGAACTGACAAATCGAAGTTCAAAATCATTTAGGGAAGAAATGGATC	1942	PR	10-JUN-1999;	99US-0138847P.
				PR	14-JUN-1999;	99US-0139119P.
Db	1748	AATGTGATGAAAATTTGCATAGTCAAAATTTTGGATTCCTTTGGATTTAAAGGAGTC	1807	PR	16-JUN-1999;	99US-0139452P.
				PR	16-JUN-1999;	99US-0139453P.
Qy	1943	TATCTGACTCTTGAAGAGGCCGT	1965	PR	17-JUN-1999;	99US-0139492P.
				PR	18-JUN-1999;	99US-0139454P.
Db	1808	TATCTCACAGTGGAGAAAGTGT	1830	PR	18-JUN-1999;	99US-0139455P.
				PR	18-JUN-1999;	99US-0139456P.
RESULT 9				PR	18-JUN-1999;	99US-0139457P.
AAC42229				PR	18-JUN-1999;	99US-0139458P.
ID AAC42229 standard; DNA; 2006 BP.				PR	18-JUN-1999;	99US-0139459P.
XX				PR	18-JUN-1999;	99US-0139460P.
AC AAC42229;				PR	18-JUN-1999;	99US-0139462P.
XX				PR	18-JUN-1999;	99US-0139463P.
DT 17-OCT-2000 (first entry)				PR	18-JUN-1999;	99US-0139750P.
XX				PR	18-JUN-1999;	99US-0139763P.
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 34765.				PR	21-JUN-1999;	99US-0139817P.
XX				PR	21-JUN-1999;	99US-0139899P.
KW Hybridisation assay; genetic mapping; gene expression control;				PR	22-JUN-1999;	99US-0140353P.
KW protein identification; signal transduction pathway; metabolic pathway;				PR	22-JUN-1999;	99US-0140354P.
KW promoter; termination sequence; ss.				PR	23-JUN-1999;	99US-0140354P.
XX				PR	24-JUN-1999;	99US-0140695P.
OS Arabidopsis thaliana.				PR	28-JUN-1999;	99US-0140823P.
PN EP1033405-A2.				PR	29-JUN-1999;	99US-0140991P.
XX				PR	30-JUN-1999;	99US-0141287P.
XX				PR	01-JUL-1999;	99US-0141842P.
XX				PR	01-JUL-1999;	99US-0142154P.
XX				PR	02-JUL-1999;	99US-0142055P.
PD 06-SEP-2000.				PR	06-JUL-1999;	99US-0142390P.
XX				PR	08-JUL-1999;	99US-0142803P.
PF 25-FEB-2000; 2000EP-00301439.				PR	09-JUL-1999;	99US-0142920P.
XX				PR	12-JUL-1999;	99US-0142977P.
XX				PR	13-JUL-1999;	99US-0143542P.
PR 25-FEB-1999; 99US-0121925P.				PR	14-JUL-1999;	99US-0143624P.
PR 05-MAR-1999; 99US-0123180P.				PR	15-JUL-1999;	99US-0144005P.
PR 09-MAR-1999; 99US-0123548P.				PR	16-JUL-1999;	99US-0144085P.
PR 23-MAR-1999; 99US-0125788P.				PR	16-JUL-1999;	99US-0144086P.
PR 25-MAR-1999; 99US-0126264P.				PR	19-JUL-1999;	99US-0144325P.
PR 29-MAR-1999; 99US-0126785P.				PR	19-JUL-1999;	99US-0144331P.
PR 01-APR-1999; 99US-0127462P.				PR	19-JUL-1999;	99US-0144332P.
PR 06-APR-1999; 99US-0128234P.				PR	19-JUL-1999;	99US-0144333P.
PR 08-APR-1999; 99US-0128714P.				PR	19-JUL-1999;	99US-0144334P.
PR 16-APR-1999; 99US-0129845P.				PR	19-JUL-1999;	99US-0144335P.
PR 19-APR-1999; 99US-0130077P.				PR	20-JUL-1999;	99US-0144352P.
PR 21-APR-1999; 99US-0130449P.				PR	20-JUL-1999;	99US-0144632P.
PR 23-APR-1999; 99US-0130510P.				PR	20-JUL-1999;	99US-0144884P.
PR 23-APR-1999; 99US-0130891P.				PR	21-JUL-1999;	99US-0144814P.
PR 28-APR-1999; 99US-0131449P.				PR	21-JUL-1999;	99US-0145086P.
PR 30-APR-1999; 99US-0132048P.				PR	21-JUL-1999;	99US-0145088P.
PR 30-APR-1999; 99US-0132407P.				PR	22-JUL-1999;	99US-0145085P.
PR 04-MAY-1999; 99US-0132484P.				PR	22-JUL-1999;	99US-0145087P.
PR 05-MAY-1999; 99US-0132485P.				PR	22-JUL-1999;	99US-0145089P.
PR 06-MAY-1999; 99US-0132486P.				PR	22-JUL-1999;	99US-0145192P.
PR 06-MAY-1999; 99US-0132487P.				PR	23-JUL-1999;	99US-0145145P.
PR 07-MAY-1999; 99US-0132863P.				PR	23-JUL-1999;	99US-0145218P.
PR 11-MAY-1999; 99US-0134256P.				PR	23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999; 99US-0134218P.				PR	26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999; 99US-0134219P.				PR	27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999; 99US-0134221P.				PR	27-JUL-1999;	99US-0145918P.
PR 14-MAY-1999; 99US-0134370P.				PR	27-JUL-1999;	99US-0145919P.
PR 18-MAY-1999; 99US-0134768P.				PR	28-JUL-1999;	99US-0145951P.
PR 19-MAY-1999; 99US-0134941P.				PR	02-AUG-1999;	99US-0146386P.
PR 20-MAY-1999; 99US-0135124P.				PR	02-AUG-1999;	99US-0146388P.
PR 21-MAY-1999; 99US-0135353P.				PR	02-AUG-1999;	99US-0146389P.
PR 24-MAY-1999; 99US-0135629P.				PR	03-AUG-1999;	99US-0147038P.
PR 25-MAY-1999; 99US-0136021P.				PR	03-AUG-1999;	99US-0147204P.
PR 27-MAY-1999; 99US-0136392P.				PR	04-AUG-1999;	99US-0147302P.
PR 28-MAY-1999; 99US-0136782P.				PR	05-AUG-1999;	99US-0147192P.
PR 01-JUN-1999; 99US-0137222P.				PR	05-AUG-1999;	99US-0147260P.
PR 03-JUN-1999; 99US-0137528P.				PR	06-AUG-1999;	99US-0147303P.
PR 04-JUN-1999; 99US-0137502P.				PR	06-AUG-1999;	99US-0147416P.
PR 07-JUN-1999; 99US-0137724P.				PR		

PR	09-AUG-1999;	99US-0147493P.	QY	107	AACAACCTTTGAGAGAGTGACCAAGTCGAGGTTCCACCGCCAAGCGGTTTTCAGAGTCT	166
PR	10-AUG-1999;	99US-0147935P.	DB	102		
PR	11-AUG-1999;	99US-0148171P.				
PR	12-AUG-1999;	99US-0148319P.				
PR	13-AUG-1999;	99US-0148341P.	QY	167	CTAAAGTACTCTTTGAAGGAGAGCTTTCTCCCTGATGACCTTTTGAGGAGAGTTCAAGAAC	226
PR	13-AUG-1999;	99US-0148565P.	DB	162		
PR	13-AUG-1999;	99US-0148684P.				
PR	16-AUG-1999;	99US-0149388P.				
PR	17-AUG-1999;	99US-0149175P.	QY	227	AAGCCAGCTTCCAAGAAAGTTTCATGCTTGCGCTTCAGTTCCTTCTTCCCATTTTTCGAATGG	286
PR	18-AUG-1999;	99US-0149426P.	DB	222		
PR	20-AUG-1999;	99US-0149722P.				
PR	20-AUG-1999;	99US-0149723P.	DB	287	GCTCCCAATACACCTTTTCAGTTCCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCT	346
PR	23-AUG-1999;	99US-0149902P.	DB	282		
PR	23-AUG-1999;	99US-0149930P.				
PR	25-AUG-1999;	99US-0150566P.	QY	347	AGCTTGGCCATTCCTCAGGGCATCAGTTATGCAAGCTGCGCAACCTCCCTCCCAATTCCT	406
PR	26-AUG-1999;	99US-0150884P.	DB	342		
PR	27-AUG-1999;	99US-0151065P.				
PR	27-AUG-1999;	99US-0151066P.	DB	407	GGTCTATCTCGAGCTTCGTGCGGATAGTTACGCGAGCTAGCAAAATTCACCACTCGTT	401
PR	27-AUG-1999;	99US-0151080P.				
PR	30-AUG-1999;	99US-0151303P.	QY	467	TTGGCAGTGGGACTGTGGGGTTGGATCGCTTCTGATGGGTTGATGTTGAGTAATGCC	526
PR	31-AUG-1999;	99US-0151438P.	DB	462		
PR	01-SEP-1999;	99US-0151930P.	QY	527	GTGTATCCCAATGAAGACCCAAAGCTTTTACCTCCACCTGGCTTTCACAGCTACATATTT	586
PR	07-SEP-1999;	99US-0152363P.	DB	522		
PR	10-SEP-1999;	99US-0153070P.	QY	587	GCTGTGTTTTTTCAGGCTGCTTGGGTTCTGTTTGTAGTGGGTTGATCGTGGATTTTCTG	646
PR	13-SEP-1999;	99US-01533758P.	DB	582		
PR	13-SEP-1999;	99US-0154018P.				
PR	16-SEP-1999;	99US-0154039P.	QY	647	TCATATGCAACATATAGGTTTCATGGAGGAGGAGCCACGGTGTGTCTGCGAGCAA	706
PR	20-SEP-1999;	99US-0154779P.	DB	642		
PR	22-SEP-1999;	99US-0155139P.				
PR	23-SEP-1999;	99US-0155486P.	QY	707	CTAAATCGATTCTTGGCTTGAGCATTTTCAACCATGGAGCTGATATCATATCAGTGATG	766
PR	24-SEP-1999;	99US-0155659P.	DB	702		
PR	28-SEP-1999;	99US-0156458P.				
PR	28-SEP-1999;	99US-0156596P.	QY	767	CGCTCTGTTTTTCAACCAAACTCATGAGTGGAGTGGGAAAGTCTGTTGTTAGGATGTGC	826
PR	04-OCT-1999;	99US-0157117P.	DB	762		
PR	06-OCT-1999;	99US-0157753P.				
PR	07-OCT-1999;	99US-0157865P.	QY	827	TTCAATTTTCTCTCTTAGCAAGATACCTTTCAGCAAAAACGACCAAGGTTTTTTTGG	886
PR	08-OCT-1999;	99US-0158029P.	DB	822		
PR	12-OCT-1999;	99US-0158369P.				
PR	13-OCT-1999;	99US-0159293P.	QY	887	GTGTGAGCAATGGCGCATTTGACGCTCGCTTATTTGGGAAGTCTCTTGGTTTATTTCACT	946
PR	13-OCT-1999;	99US-0159294P.	DB	882		
PR	14-OCT-1999;	99US-0159329P.				
PR	14-OCT-1999;	99US-0159330P.	QY	947	CACGCCGAGAGCAGCGTGTGGAAGTGTAGGAGAACTGGAAGAGGTTTGAATCCACCA	1006
PR	14-OCT-1999;	99US-0159637P.	DB	942	AGAGCCGAGCGTCACGGAATCAGCGTCATCGGGAATACCAGAAAGTTTGAATCCACCG	1001
PR	14-OCT-1999;	99US-0159638P.				
PR	18-OCT-1999;	99US-0159584P.	QY	1007	TCACTCAGAAATCTGGTATTTGTGCGCTTACATGACTACAGCTGTCAAACTGGCAT	1066
PR	21-OCT-1999;	99US-0160741P.	DB	1002		
PR	21-OCT-1999;	99US-0160767P.				
PR	21-OCT-1999;	99US-0160768P.	QY	1067	GTGCTTGGCATCATATCACTTCGGGAAGGAATAGCAGTAGGAGAGCTTTGCAATGTAT	1126
PR	21-OCT-1999;	99US-0160814P.	DB	1062		
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.	QY	1127	AAAAATTAACAATTTGATGGCAACAAAGATGATAGCTATTTGGGACCACTGAACGTAGTT	1186
PR	22-OCT-1999;	99US-0160981P.	DB	1122	AGAACTACCACTAGTAGTGAAACAAAGAGATGATGCCCATTTGGTCTGATGAACGTAGTA	1181
PR	23-OCT-1999;	99US-0161404P.				
PR	25-OCT-1999;	99US-0161405P.	QY	1187	GGTCTTTTCACTCTTTGTGCTACCTCAACAGGACCAATTTTCGGGTTCCGGCTGTGAACAT	1246
PR	25-OCT-1999;	99US-0161406P.				
PR	26-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.	DB			
PR	26-OCT-1999;	99US-0161361P.				
PR	28-OCT-1999;	99US-0161920P.	QY			
PR	28-OCT-1999;	99US-0161922P.	DB			
PR	28-OCT-1999;	99US-0161993P.				
PR	29-OCT-1999;	99US-0162142P.				

Query Match 27.0%; Score 662.4; DB 3; Length 2006;
 Best Local Similarity 60.1%; Pred. No. 6.8e-153;
 Matches 1120; Conservative 0; Mismatches 741; Indels 3; Gaps 1;

Db 1182 GGCTCTGCCACATCTTGCTACGTCAACCCGGAGCATCTCTAGATCAGCGGTCAACAAC 1241
Qy 1247 AACGCTGGATGCAAGACAGCTTCCACATTAATATGTCACCTGACGTAATGTTGACA 1306
Db 1242 AACCGGGAGCTAAACCGCAGTTTCAACATTTGTTATGTCGCTACTGTTATGGTTACG 1301
Qy 1307 TTCTTATTCCTGACACCCCTGTTCCATTAACATCCCTCGTGGTGTCTATCAGCTATTATC 1366
Db 1302 CTCTCTCTATGCGCTTTTCGATACACTCCCAATGTGCTCTCGGTGCCATCAT 1361
Qy 1367 GTATCTGAATGCTTGGACATCAGATGATTAAGACAGCCCATCATTTTAAAGTTGAC 1426
Db 1362 GTACCGCGCTCATTTGCTCTATCGACCTTCCCGCGCTGTACATATGGAAGATCGAT 1421
Qy 1427 AATTTGACTTTGCTGTGTCATGATGCGATACATTTGGCGTGTCTTGGCAGTGTGAA 1486
Db 1422 AATTTGATTTCTTGTGTATGCTTTGGCGGTCTTTTGGTGTCTTTTCTATCCGTTCAA 1481
Qy 1487 ATTGGCTTAGTCATAGCTATTGTAATATCTGTACTTTCGGGTACTTCTATTATTGCAAG 1546
Db 1482 AACGGCTAGCCATTGGGTGGGCTATCGTTGTTCAAGATATTGATGCAAGTAAACAAG 1541
Qy 1547 CCAAGACATCTGTTTGGGCAACATTCCAAATCTGTGATATACCGAAATGTTGACAC 1606
Db 1542 CCGAAAATGGTTATAATGGGTAAATTTCTTGGAAACGATATATACCGAGATCTTCATCAT 1601
Qy 1607 TATCAAAATGCRAAACATGTTCTGGAATGCTAAATTTCTAGAGATTGATGCCCAATTTAC 1666
Db 1602 TACAAAAGAGCAAAAGGATCCCGGATTTCTTGTTTTAAAGCATCGAATCTCTCTGTCAT 1661
Qy 1667 TTTGCCAATGCCAGCTATTTTAAGAGAAAGGATCACAAAGTGGATTGATGAAGAAGAA 1726
Db 1662 TTGCGCAATTTAACTACCTACTGTAAGAAACATCTCGTTGGATTGAAGATCCGAAGA 1721
Qy 1727 AGAATTAAGCTACAGGGAGACTAGTTTGCAGTATGTTTAAATGATATGAGTGTCTGT 1786
Db 1722 GAGGAAGCTCAAGAGAAGCATTTAGCCTACAGTTCTTGATTTCTTGAATGTGAGCGGTG 1781
Qy 1787 GGAACATTTGATACAGTGTGAATAAGTATGCTTGAAGAGTGAAGAGATTTACAGAGAGA 1846
Db 1782 AGCGGTGTAGACAAACCGAGTGTCTTTTAAAGAACTTTAAAGAAACACCCGCCAAG 1841
Qy 1847 AGAGACTACAGCTTGTGTTTGGTCAATCTCTGAAGTGAAGTGAAGAACTGAACAAA 1906
Db 1842 AAGGACATCGAGCTTGTGTTGTGAACCTCTAAGCGAAGTGTGGAAGCTTCAAGA 1901
Qy 1907 TCGAAGTTCCAAATCAATTTAGGGAAG---AAATGGATCTATCTGACTGTTGAAGAGCC 1963
Db 1902 GCTGACGAACAAAAGAGTTTCATGAGGCCCGAGTTTCTCTTTAAACCGTCGTGAGGCC 1961
Qy 1964 GTTG 1967
Db 1962 GTTG 1965

RESULT 10

ADA68554
ID ADA68554 standard; DNA; 1896 BP.

XX ADA68554;

XX AC
XX DT
XX 20-NOV-2003 (first entry)

XX Arabidopsis thaliana gene, SEQ ID 528.

DE Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX OS
XX PN
XX WO2003000898-A1.

XX PD
XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.
XX 22-JUN-2001; 2001WO-IB001105.
PR (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
PS Claim 6; SEQ ID NO 528; 899pp; English.
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX Sequence 1896 BP; 483 A; 465 C; 426 G; 522 T; 0 U; 0 Other;

Query Match 27.0%; Score 660.2; DB 8; Length 1896;

Best Local Similarity 60.2%; Pred. No. 2.3e-152; Indels 3; Gaps 1;

Matches 1113; Conservative 0; Mismatches 733;

Qy 122 GTGCACCAAGTCGAGGTTCCACCGCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTG 181
Db 7 GTGCACCAAGTGTGTTGCTCCGCGCATTAAGACACGGTGGCAAGCTAAAGACAAAAC 66
Qy 182 AAGGAGACTTTCTTCCCTGATGACCTTTGAGGCGAGTTCAAGAACAGCAGCTTCCAG 241
Db 67 AAGGAGACTTTCTTCCGAGCAGTCCCTTAAGACAGTTTCAGAGGACCAACCAACCGTACC 126
Qy 242 AAGTTTCATGCTGGGCTTCAGTCTTCTTCCCATTTTCGAATGGCTCCCAATACACC 301
Db 127 AAGCTCATACGCGCGCTCAATAGATTTTCAATCTCCATGGTGCTCAGGTACAGC 186
Qy 302 TTTCAAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGTAGCTTGGCCATTCT 361
Db 187 TTCAGCTACTCAAATCCGAGCTGCTTTCCAGTCTTACCATCGTAGTTAGCTATTCT 246
Qy 362 CAGGGCATCAGTTATGCCAAGCTCCCAACCTCCCTCCATTTCTTGGACTATTTCCGAGC 421
Db 247 CAGGGGATTTAGTTACGGAAGCTAGCAAAATTTGCCCAACCAATCGTTGGTCTATTCTCGAGC 306
Qy 422 TTTATACCAACCATTTGATTTATCGGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACT 481
Db 307 TTCGTGCCACCGTTGTTATGCGGTCTTGGGAAGCTCAAGAGATCTAGCAGTGGGACCA 366
Qy 482 GTGGCGTTGGATCGCTTCTGATGGGTTTCGATGTTGAGTAAATGCGGTTGATCCCAATGAA 541
Db 367 GTCTCCATAGCGTCGTTAATCTTAGGATCCATGCTAAGGCAACAAGTATCTCCCGTAGAC 426
Qy 542 GACCCAAAGCTTTACTCCACCTGCTTTCAGAGTACATATTATTTGCTGGTGTCTTTTCAG 601
Db 427 GATCTGTCTCTTTCTACAGCTAGCTTCTCTTCTACCTTCTTCTGCTGGTCTCTTTCAA 486
Qy 602 GCTGCTTGGGTCTGTTTGGGTTGGGTGATGCTGGGATTTTCTGTCATCATCAACCAATA 661
Db 487 GCCTCTCTTGGNAATCTCAGGCTGGGATTTAATAGACTTTCTATCAAAAGGAGCTTA 546
Qy 662 ATAGGGTTCAATGGGAGGAGAGCCACGGTGGTGTCTCTGACGAACTAAATATCGATTCTT 721

Db 547 ATAGGGTTATGGGTGAGCAGCCATAATTTGTAATCACTCAACAGCTAAAGGGTCTGCTT 606
Qy 722 GGCTTTGAGCATTTCAACCATGGAGCTGATATCATATCACTAGTATGCGCTCTGTTTTCAAC 781
Db 607 GGGATRAACTATTTCACAAAGCATATGAGTGTAGTCCCTCTCTCTCTCTGTTTTCCAA 666
Qy 782 CAAACTCATGAGTGGAGGTGGAAAAGTCTGTGTGTAGGATGTGTCTTTCAATTTCTCTC 841
Db 667 CACACCAACGAGTGGTCAATGGCAACAATTTGATGGAGTGTGCTTTCTGCTCTCTT 726
Qy 842 CTTTAGCACAAGATACTTTCACAAAAACGACCAAGGTTTTTTTTGGGTGTCAGCAATGGCG 901
Db 727 CTCTGCACAGTCACCTCAGCATGAAGACCGAAGCTGTTTTGGGTCTCAGCGGAGCA 786
Qy 902 CCATTGACGTCGGTTATTTGGGAAGTCTCTTGGTTTATTTCACTCACGCCGAGAAACAC 961
Db 787 CCACCTCTTTCCGTTATCGTCTCTACACTTCTTGTCTTCTTTCAGAGCGGAGCTCAC 846
Qy 962 GGTGTTGAAGTATAGAGAACTGAAGAAGGGTTTGAATCCACATCACTCACAAAATCTG 1021
Db 847 GGAATCAGCGTCAATCGGAAATTTACCAAGAGGTTTGAATCCACCGTCTTGGAAACATGCTT 906
Qy 1022 GTATTTGTGTCGCTTACATGACTACAGCTGTCAAAACTGCGCATTTGCTGTGCGCATATA 1081
Db 907 CAGTTTCACGGTAGTCATCTCGCACTGTCGCCAAAACCGGACTCGTTCACCGGAATCGTC 966
Qy 1082 TCACCTTCGGAAGGAATPAGCAGTAGGAAGAGCTTTTGCATGTATATAAAATTTACAATATT 1141
Db 967 TCCTTCACGGAAGGAATCGCAGTGGGAGAACATTTGCAGCGCTAAGAACTACACCGTA 1026
Qy 1142 GATGGCAACAAAGAGATGATAGTATTTGGGACCATGAAGCTAGTGTGTTCTTTCACTCT 1201
Db 1027 GATGGAACAAAGAGATGATCGCCATTTGTCGTCACCTGTTATGTTAGCTTCTCTCTTAATG 1206
Qy 1202 TGCTACTCTCAACAGACACATTTTCGCGTTCGCTGTGAATCTAATACGCTGATCGAAG 1261
Db 1087 TGCTAGTCTCAACCGGAGCAATTTCTTAGATCAGCGGTCAACAAACACGCGGAGGTAA 1146
Qy 1262 ACAGCAGCTTCCAAACATTAATATGTCACTTTCGAGTAATGTTGACATTTGTTATTTCTTGACA 1321
Db 1147 ACCGAGTTTCAAAACATTTATGTCGGTCACCTGTTATGTTAGCTTCTCTCTTAATG 1206
Qy 1322 CCCTTGTTCATTAACATCCCTCGGTGGTGTATCAGCTATTAATCGTATCTGCAATGCTT 1381
Db 1207 CCGCTTTTCGAATACATCTCCAAATGTGTCCTCGGTGCCATCATTTGTACCGCGGTCAIT 1266
Qy 1382 GGACTCATAGATTATGACAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTGTG 1441
Db 1267 GGTCTCATCGACCTTCCCGCGGCTGTACATATGGAAGATCGATAAATTTGATTTCTTGT 1326
Qy 1442 GTGTGATGAGTCATACATTTGGTGGTCTTTTGGCAGTGTTCGAAATTTGGCTTAGTCATA 1501
Db 1327 GTGATGCTTTGCGGTCTTTTGTGTCAATTTCTTCTATCCGTTCAAAACGCTTAGCCATT 1386
Qy 1502 GCTATTGTAATATCTGTACTTCGGGTACTTCTATTTATTGCAAGGCCAAGCAATTCGTT 1561
Db 1387 GCGGTGGGGTATCGTTGTTTCAAGATATTGATGCAAGTAAACAAAGGCCGAAATGGTTATA 1446
Qy 1562 TTGGGCAACATTCCAAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATGCAAA 1621
Db 1447 ATGGGTAATATTCCTGGAACGGAATATATACCGAGATCTTCATCATTAACAAAGAACACAA 1506
Qy 1622 CATGTTCTCTGGAATGCTAAATTTCTAGAGATTGATGACCACTTTTACCTTTTGGCAATGCGAGC 1681
Db 1507 AGATCCCGGATTTCTTGTGTTTAAAGCATCGAATCTCTGTCAATTTGCGCAATTTCTAAC 1566
Qy 1682 TATTTAAGAGAAAGGATCAACAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACA 1741
Db 1567 TACCTCACTGAAAGAACATCTCGTTGGATTGAAGAAATGCGAAGAGGAAGCTCAAGAG 1626
Qy 1742 GGGGAGACTGTTTGCAGTATGTTATAATTGATATGAGTCTCTTGGAAACATTTGATACA 1801

Db 1627 AAGCAATCTAGCTACAGTCTTCTTGATCTTTGAAATGTCAGCCGTCAGCGGTAGACACA 1686
Qy 1802 AGTGGAAATAGTATGCTTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGCTACAGCTT 1861
Db 1687 AACGGAGTGTCTTTTAAAGGAACTTAAGAAAAACAACGCCCAAGGAAGGACATCGAGCTT 1746
Qy 1862 GTTTTGGTCAATCTCTTAAGTGAAGTGAAGAAAACCTGAACAAATCGAAGTTCCAAAT 1921
Db 1747 GTGTTTGTGAACCTCTTAAGCGAAGTGTGGAGAAGCTTTCAAAGAGCTGACGACAAAA 1806
Qy 1922 CATTTAGGGAAG--AAATGGATCTATCTGACTGTTGTAAGAGGCCGTTG 1967
Db 1807 GAGTTTCATGAGGCCCGAGTTTCTTCTTTAAACCTGCTGAGGCCGTTG 1855

RESULT 11
AAZ50481
ID AAZ50481 standard; cDNA; 2279 BP.
XX AAZ50481;
XX
DT 23-MAY-2000 (first entry)
XX
XX Corn sulphate permease-1 cDNA clone.
XX
XX Sulphate Permease; sulphate assimilation protein; corn; contig; probe;
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
KW antibody; screen; ss.
XX
XX Zea mays.
XX
FH Key Location/Qualifiers
CDS 3..2069
FT /*tag= a
FT /product= "Corn sulphate permease-1"
FT /notes= "Derived from a contig composed of clones
FT cbn10.pk0062.b10, ccoin.pk081.b21, ccoin.pk092.12,
FT csc1c.pk005.j3, p0004.cb1ej58r, p0089.csdch19r,
FT p0094.csssg12r, p0121.cb1rmx30r and p0128.cb1c209r"
FT
XX
XX WO200004154-A2.
XX
XX
XX 27-JAN-2000.
XX
XX 13-JUL-1999; 99WO-US015810.
XX
XX 14-JUL-1998; 98US-0092833P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Falco SC, Thorpe CJ;
XX
XX WPI: 2000-195025/17.
XX P-PSDB; AAY44935.
XX
XX Nucleic acid fragments encoding sulfate assimilation proteins in plants
XX and seeds useful as probes for isolating cDNAs and genes encoding
XX homologous proteins, in producing transgenic plants.
XX
XX Claim 3; Page 38-39; 79pp; English.
XX
XX The present sequence is a cDNA clone encoding corn sulphate permease, a
XX sulphate assimilation protein. This sequence is from a contig composed of
XX clones cbn10.pk0062.b10, ccoin.pk081.b21, ccoin.pk092.12, csc1c.pk005.j3,
XX p0004.cb1ej58r, p0089.csdch19r, p0094.csssg12r, p0121.cb1rmx30r and
XX p0128.cb1c209r from cbn10, ccoin, csc1c, p0004, p0089, p0121 and
XX p0128 libraries, respectively. This sequence is used as a probe to
XX isolate other plant sulphate assimilation proteins, for genetic and
XX physical mapping of related genes and as markers of traits linked to the
XX gene. This is useful for plant breeding and to construct chimeric genes,
XX used to create transgenic plants with altered levels of sulphate
XX permease. The sulphate permease peptides are useful for producing
XX antibodies, that are used to screen and isolate cDNA clones

XX	Sequence	2279 BP; 512 A; 606 C; 629 G; 532 T; 0 U; 0 Other;	
SQ	Query Match	26.8%; Score 655.4; DB 3; Length 2279;	
	Best Local Similarity	60.4%; Pred. No. 3.8e-151;	
	Matches 1120; Conservative	0; Mismatches 726; Indels 9; Gaps 2;	
QY	121	AGTGCAACCAAGTCAGAGTTCCACCGCCACAGCCGGTTTTTCAAGTCTCTAAAGTACTCTTT	180
Db	188	AGTGCAACAGATGGCGCCGCCACCGCGCGGAGCACGGCGAGCAAGATGAAGTGAGGTT	247
QY	181	GAAGGAGACTTCTTCCCTGATGACCCCTTTGAGGCAGTTCAAGAACAAAGCCAGCTTCCAA	240
Db	248	GAAGGAGACCTTCTTCCCGCAGACCCGCTTCGGGGGTTCAAGGGGCGACGCGCGGGGAC	307
QY	241	GAAGTTTCATGCTTGGCTTCAGTTCCTTCTTCCCATTTTCGAATGGGCTCCCAATACAC	300
Db	308	GCAGTGGCTCATGGCGGTTCAGGTACCTCTTCCCATCTCTGGAGTGGGTCGGAGCTACTC	367
QY	301	CTTTCAAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTC	360
Db	368	CTTGTGCTCTTCAAGTCCGACCTCGTCGGGGCTCACCATTTGCCAGCTCGCCATTC	427
QY	361	TCAGGCGATCAGTTATGCCAAGCTCGCCAACTCTCCCAATTCCTTGGACTATATTCGAG	420
Db	428	TCAGGCGATTAGCTACGCCAAGCTGCCAAGCTTGCCCTCCCATAAATCGGGCTGTATTCGAG	487
QY	421	CTTTATACCAACCATTCATTTATCGCATGATGGTAGCTCGAGGGATTTGGCAGTGGGAC	480
Db	488	CTTGTGTCGGCCGATGGGTACCGGTGTCTGGGGAGCTCCCGTAGCTGGCGGTGGGCC	547
QY	481	TGTGGCGGTGGATCCGCTTCTGATGGGTTCCGATTTGAGTAATGCCGTTGATCCCAATGA	540
Db	548	GGTGTGATCTCGTCTGATCATGGGTCCATGCTGGCGCAGCGCTGAGCCCCACTGC	607
QY	541	AGACCCAAAGCTTTACTCCACTGCTTTCAAGCTACATTAATTTGCTGGTGTTTTCA	600
Db	608	GGAGCCGAGCGTGTCTCGCAGCTGGCTTCACTCCACCTGTTCGCGGGGTGGTGCA	667
QY	601	GGCTGCTGGGTCTGTTAGTTGGGTGATCGTGGATTTCTGTACATGCAACCAT	660
Db	668	GGCTTCCCTGGGGATCCTCAGGCTCGGCTTCGTCATCGACTTCCCTGTCCAAGCGCAGCT	727
QY	661	AATAGGGTTTCATGGGAGGAGCACCAGCGTGGTGTCTGCAGCACTAAATTCGATTCT	720
Db	728	GGTGGGGTTTCATGGCGCGCGCCCATCATCGTGGCGCTGCAGCACTCAAGGGCTGCT	787
QY	721	TGGCTTTGAGCATTTCAACCATGGAGCTGATATCATATCAGTGATCGCTCTGTTTTCAC	780
Db	788	GGGCATCGTCCACTTCACCACCGAGATGGCATCGTCCCAGTCAATGGCCTCGTCTTTCA	847
QY	781	CCAAACTCATGTGAGTGGAGTGGAAAGTCTGTGTTAGGATGTGTTTCATTTTCTTCT	840
Db	848	CCACACGAGGAGTGTGTCGTCGACAGCATCTCATGGGCGTCTGCTTCTCGTCTTCT	907
QY	841	CCTTAGCAACAAGTACTTCAGCAAAAACGACCAAGGTTTTTTGGGTGTGACGAATGGC	900
Db	908	GCTGTCCGCGAGGCATGTGGCATCAGATGGCCAAAGCTTTCTGGGTTTCGGCGTGGC	967
QY	901	GCAATTGACGTCGGTTATATGGGAAGTCTCTTGGTTTATTTTCACTCAAGCGGAGAGCA	960
Db	968	GCCCCGGCATCGGTCAACCATCTCGACGCTGTGTTTTCCTCTTCAAAAGCTCAGAAACA	1027
QY	961	CGGTGTTGAAGTATAGGAGACTGAAGAAGGTTTGAATCCACCATCACTCACAAATCT	1020
Db	1028	TGGCATCAGCATCAITGGGAGCTCAAGTGGCGCTGGAATCGCCCTCGTGGGACAAGCT	1087
QY	1021	GGTATTTGTTCGCCCTTACATCACTACAGCTGTCAAAACTGGCATTTGCTGGCATCAT	1080
Db	1088	CCTGTTTGACCGGCGTATTAGGCGCTCACCATGAAGACTGGCTTGTCAACCGGAATCAT	1147
QY	1081	ATCATTGCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAATTTACAAAT	1140

Db	1148	CTCACTGACGGAAGGAATAGCGGTTGGTAGAAACATTTGCCTCACTCAAGGACTACCAGAT	1207
QY	1141	TGATGCAACAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTCACTC	1200
Db	1208	AGATGAAACAAGGAGATGATGGCCATAGGGTTGATGAATGTTTGGGTCTTGCAATC	1267
QY	1201	TTGCTACCTTCACAAACAGGACCAATTTTCGGCTTCGGCTGTGAACACTATAACGCTGGAT	1260
Db	1268	ATGCTACGTAAACAACAGGTGCGTTCTCCGCTCTGCTGTAAACCAACAACGCGGCTG	1327
QY	1261	GACAGAGCTTCCAACATTAATATGCTACTTCGAGTAATGTTGACATTTGTTATTCCTGAC	1320
Db	1328	GACTGCCATGTCCAACGCTGATCATGGCGTGAATGATGATGTCAGCTCTGTTCTCAT	1387
QY	1321	ACCTTGTTCATACACTCCCTCGTGGTATACAGCTATATGCTATATGCTATGCAATGCT	1380
Db	1388	GCACTGTTTCGTGTACACACCAACGCTTGTCTCGGAGCGATCATCGCCGCGGTGAT	1447
QY	1381	TGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTGT	1440
Db	1448	CGGCTGATCGATTTCCCGCGGTGTACACATCTGGAAGATGGACAAGATGGAATTTCT	1507
QY	1441	GGTGTGATGATGATACATTTGGCGTGGTCTTTTGGCAGTGTGAAATTTGGCTTAGTCA	1500
Db	1508	GGTGTGCGTTTCGGGTTTTCGGCGGTTCATCTCTCAGTCCAAGAAAGGCTTGGCAT	1567
QY	1501	AGCTATTTGTAATCTGTACTTTCGGGTACTTCTATTTATTGCAAGCCCAAGACATTCGT	1560
Db	1568	ACGGTGGTATATCTATATTTAGGGTGTGATGAGATCAACAGGCCGAGATGATGTT	1627
QY	1561	TTTGGGCAACATTTCCAAATTTGTGATATACCGAAATGTTTGAGCACTATCAAAATGCAAA	1620
Db	1628	TCAAGGGAACATCAAGGGGACTGATATTTACAGAGACCTGTCATCACTACAAGGAGCCCA	1687
QY	1621	ACATGTTCTGGAATGCTAAITTCAGAGATTGATGACCAATTTACTTTGCCAATGCCAG	1680
Db	1688	AAGAGTTTCTGGGTTCTTGATCTTGGCCATTGAAGCACCGATAAACTTCGCCAACTCCA	1747
QY	1681	CTATTTAAGAGAAAGGATCACAAAGTGGATTGATGAAGAAAGAAAGAAATTTAAAGCTAC	1740
Db	1748	CTACCTGAATGAAGGATTTAAAGATGGAT-----AGAGGAAGAAATCTTTTGAACAGGA	1801
QY	1741	AGGGGAGACTAGTTTGCAGTATGTTATAATGATATGATGCTGTTTGGAAACATTTGATAC	1800
Db	1802	TAAACATACTGAACCTCCATTTTCATAATCTTGATCTGTCAAGCTGTTCTGCAATTGACAC	1861
QY	1801	AAGTGAATAAGTATGCTTGAAGAGGTGAAGAATTACAGAGAGAGAGAGCTACAGCT	1860
Db	1862	AAGTGGCATAGCGTTCTCCTCAITTGACATAAAGAAATCAATAGAGAAACGTTGGAGCT	1921
QY	1861	TGTTTTTGGTCAATCCTGTAAGTGAAGTGAAGAAACCTGAAACAAATCGA---AGTTCCA	1917
Db	1922	TGTGCTGTGTAATCACTCAACTGGAGAGTCAATGGAGAAATACAACTGCAACAGAGCTGA	1981
QY	1918	AAATCTTTAGGGAAGAAATGGAATCTATCTGATCTGTTGAAGAGGCGGTTTGGAGCA	1972
Db	1982	AAACTATTTTAGGCCAGATTGCTTGTATCTGACCCTGCGGAAGCAATCGCTTCA	2036

RESULT 12

ADA70466

ID ADA70466 standard; DNA; 1989 BP.

XX ADA70466;

XX ADA70466;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 3789.

XX Plant; bacterial infection; fungal infection; viral infection; rice;

XX gene; db.

XX Oryza sativa.

Db 1669 TATGTTAAAGACAGATTCCTGAGGTGGCTGAGGAGCAGGAGAAAGAGTTCAGCAGGAACAG 1728
QY 1742 GGGGAGACTAGTTTGCAGTATGTTATAATTGATATGAGTGCTCTTGGAAACATTGTGATACA 1801
Db 1729 AAGCTACAGAAACTGAGTTTCTGATTTGTTGAGCTATCTCCGGTAATTGACATTGTATACA 1788
QY 1802 AGTGGAAATAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTTACAGCTT 1861
Db 1789 AGTGGAAATCCATGCTTGGAGGATCTATTATAGAGCACTTGAAAGCGCAAAATCCAGCTG 1848
QY 1862 GTTTTGGTCAATCCTGTAAGTGAAGTGAAGAAGAACTGAACAAATCGAAGTTCCAAAT 1921
Db 1849 ATTCTAGCCAATCCGGGGCAGCTGTGATCTCTGAAGCTCCGATCAGCGAAATTCACGGAT 1908
QY 1922 CATTTAGGGAAGAAATGGATCTATCTGACTGTGTTGAAGAGGCGCT 1965
Db 1909 CTCATCGGTGAAGACAAGATATTCCTGACTGTGCGGAGCGAGT 1952

RESULT 13

ADA69770
ID ADA69770 standard; DNA; 1947 BP.

XX AC ADA69770;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 3093.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.

PS Claim 6; SEQ ID NO 3093; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

XX Sequence 1947 BP; 515 A; 481 C; 472 G; 479 T; 0 U; 0 Other;

XX Query Match 26.4%; Score 647.2; DB 8; Length 1947;

XX Best Local Similarity 60.3%; Pred. No. 3.8e-149;

XX Matches 1110; Conservative 0; Mismatches 718; Indels 12; Gaps 2;

QY 126 AACAAAGTCGAGGTTCACCGCCACACGCGCTTTTCAAGTCTCTAAAGTACTCTTTGAAGG 185
Db 83 ACAATGTCGGCGCGCGCGCGAAGAACTCTCTGGCGAGTTCGCGCGCAGGTGAAGG 142
QY 186 AGACTTTCTTCCCTGATGACCTTTTGGAGCAGTTCGAAGAACAGCCAGCTTCCAAGAGT 245
Db 143 AGACGCTTCTTCTCCAGCAGCGGATCGCGCGGTACAAGGACCAGCGAGGTCCAGGAAGC 202
QY 246 TCATGCTTGGCCTTCAGTTCCTTCTTCCCATTTTGGAAATGGGCTCCCAATACACCTTTC 305
Db 203 TATGGCTGCCCTTGGAGCACGCTTCTCCGGTGTTCGAATGGGGCAGACAATACACCTCG 262
QY 306 AGTTCCTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAGG 365
Db 263 CCAAGTTCAAGGGCGACCTCATTTGCCGCGCTCACCTTGGCAGCCTCGTCATACCTCAG 322
QY 366 GCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATTTCTGGACTATATTCGAGTTTA 425
Db 323 ACATCGGCTACGCGAAGCTTGCTAACTGCCACAGAGATTGGGCTGCACAGTAGCTTCG 382
QY 426 TACCACCAATGATTTATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAGCTGTGG 485
Db 383 TCCCGCCATTGATATACGCTCTGATGGGCACCTCAAGGGAGCTAGCAATGGGTCCAGTGG 442
QY 486 CGTTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTTGTATCCCAATGAAGACC 545
Db 443 CGTCACTCACTGCTGCTTGGTACTCTCTCCAGGAGGAGATTGACTCAAGAAGAAC 502
QY 546 CAAAGCTTTACCTCCACCTGGCTTTCAAGCTACATTAATTTGCTGGTGTGTTTTTCAGGCTG 605
Db 503 CGCTAGATTACAGAGCGGCTCGCTTTCACAGCGACCTTCTTTGACGAGGTTCACACAGCGG 562
QY 606 CCTTGGCTCTGTTAGTTGGGGTTCATGCTGGATTTCTGTCACATGCACCAATCAATATAG 665
Db 563 CGCTGGGTTTCTGAGGCTAGGTTTCATATAGCGTTCTGTCATGCTGCCATCATCG 622
QY 666 GGTTCATGGGAGGAGCAGCACGCTGGTGTCTGCAGCAACTAAATCGATTTCTTGGCC 725
Db 623 GATTCATGCGCGCGCGCCCATCACCATTTGCTCTTCCAGCAGCTTAAAGGCTTCTTTGGAA 682
QY 726 TTGAGCATTTACCCATGGAGCTGATATCATATCAGTATGCGCTCTGTTTTTCAACCCAAA 785
Db 683 TTGCAAACTTCAACCAAGAGACTGACATCATCTCGGTCAATGAATCAGTCTCGGGAAATG 742
QY 786 CTCATGAGTGGAGGTGGGAAAGTGTGTTTAGGATGTCTTCATTTTCTTCTCTCTTA 845
Db 743 TTCACATGG-----GGCAAGTATTTTGTACCTCTCCAACTTCTTGAT-CTCT 790
QY 846 GCACAAGATACCTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGCCAT 905
Db 791 TCTCTGAAAAAGGGGGCAAGAGAACAGAAAGCTCTTCTGGGTCCAGCAATTCACCCAC 850
QY 906 TGACGTCGTTATATTTGGGAAGTCTCTTGGTTTATTTTCACTCAGCGGAGAGACGGTG 965
Db 851 TCAATTCGCTGATCATTTTCAACTTTTTCGCTACATCATCTGCTGTGACAAACAAGGCG 910
QY 966 TTGAAGTGATAGGAGAACTGAAGAAAGGTTTCAATCCACCATCACTCAAAAATCTGGTAT 1025
Db 911 TCGCAATGTAATAAAGCTCAAGAAAGCATCAATCCACCTTCAGCTAGCCTGATATTTT 970
QY 1026 TTGTGTCGCTTTACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATCATATCAC 1085
Db 971 TCACTGGCCCACTACTTGTCTCAAAGGATTCAAAATTTGGAGTAGTAGCTGGAATGATAAGCC 1030
QY 1086 TTGCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATATAAATTTACATATTTATG 1145
Db 1031 TTACGGAAGCGATTTCAGTTGGGAAGACATTTTGTGGATTGAACGATTAACAGATAGATG 1090
QY 1146 GCAACAAAAGAGATGATAGCTATTGGGACCATGAACGTTAGTTTCTTCTTCTTCTTGTCT 1205
Db 1091 GGAACAAAAGAAATGTTGGCTCTAGGAACCATGAATGTGGTTCGTTCAATGACGCTTGTCT 1150
QY 1206 ACCTCACAAACAGGACCAATTTTTCGCGTTTCGGCTGTGAACCTATTAACGCTGGATGCAAGACAG 1265

Db 1151 ATATAGCCACAGGTGGTTTTCACGATCAGAGTCAATTGCATGGCTGGAGGTAACAC 1210
Qy 1266 CAGCTTCCAACTATATATGCTCAGTTCAGTAATGTTGACATTTGTTATCTCGACACCCCT 1325
Db 1211 CAATGTCCTAATATGTTATGTAACCTAGTATGTTGCTTGCACCTGTTGGATCACTCCAT 1270
Qy 1326 TGTTCATATACATCCCTCGTGGTGTATCAGCTATATCGTATATCGTAATCGTAATCGTGGAC 1385
Db 1271 TGTTCAGTACACGCCAATGCCACCATTTCTCCATCATCATATCAGCAGTGTGGCC 1330
Qy 1386 TCATAGATTATGAAGCAGCATCATCTATTTAAGTTGACAAATTTGACTTTGTGGTGT 1445
Db 1331 TATTTGACTTTGATCAGCTACTTATCTGGAAGTTGATAAGTTGGACTTCATGGCGT 1390
Qy 1446 GCATGAGTGCATACATTTGGCGTGTCTTTGGCGATGTTGAAATTTGGCTTAGTCAATGCTA 1505
Db 1391 GCTTGGGGGCAATTCCTTGGAGTAATATTTTCATCTGGAGTATGGCTGCTCATTTGGCG 1450
Qy 1506 TTGTAATATCTGACTTCGGGTACTTCTATTTATTCAGAGGCCAAGACATTCGTTTGG 1565
Db 1451 TTGTAATATCTGACTTCAATCAAAAGTTCTGCTCCATGTAAACACGCCCAAGCAGCTTTACTTG 1510
Qy 1566 GCAACATTCCAAATTCCTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATG 1625
Db 1511 GCAACCTTCCAGAACGATTTATCTATAGAAATGTTGAACAATATCCAGAACTTACCAAGG 1570
Qy 1626 TTCTTGAATGCTAAATTTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAGCTATT 1685
Db 1571 TGCAGGAGTCTAATTTGAAGTGGACTCAGCAATATATCTTCAAACTCCAAATATG 1630
Qy 1686 TAGAGAAAGGATCACAAGGTGATGATGAAGAAGAAAGAAATTAAGCTTACAGGGG 1745
Db 1631 TTAAGAAAGAAATGCTGAGATGCTGAGATGAGGAAGAAACATCAAAAGGAACAGAGT 1690
Qy 1746 AGACTAGTTTGCAGTATGTTAATATGATATGATAGTCTGTTGAAACATTGATACAGTG 1805
Db 1691 TACCAAAATTGAGTTCTGATTTGACCTATCTCTCTGTAATGATATTTGACACAAGTG 1750
Qy 1806 GAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGCTACAGCTTGTGTT 1865
Db 1751 GAATCCATGCTTCAAAAGAGTTGTTGAGGACACTTGAAAAGCCGAGATTCAGCTGATTT 1810
Qy 1866 TGCTCAATCTGTAAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAATCAAT 1925
Db 1811 TCCCCAATCTCGGGCGGTGTGATCCAAAGCTCCGGTCAGCAAAATTCACAGAGCTCA 1870
Qy 1926 TAGGGAAGAAATGATCTATCTGACTGTTGAAGAGGCCGT 1965
Db 1871 TTGGTGAAGAAAGATATGCTCAGAGTTGGTGACGCCGT 1910

RESULT 14
AAZ50490
ID AAZ50490 standard; cDNA; 2311 BP.
XX
AC AAZ50490;
XX
DT 23-MAY-2000 (first entry)
XX
DE wheat sulphate permease-1 cDNA clone.
XX
KW Sulphate Permease; sulphate assimilation protein; wheat; probe; mapping;
KW marker; plant breeding; chimeric gene; transgenic plant; antibody;
KW screen; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 100..2070
FT /tag= a
FT /product= "wheat sulphate permease-1"
FT /note= "Derived from clone wki.pk0028.e1"

XX
PN WO200004154-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015810.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Falco SC, Thorpe CJ;
XX
DR WPI; 2000-195025/17.
XX
DR P-PSDB; AAY44944.
XX
CC Nucleic acid fragments encoding sulfate assimilation proteins in plants
CC and seeds useful as probes for isolating cDNAs and genes encoding
CC homologous proteins, in producing transgenic plants.
XX
CC Claim 3; Page 55-56; 79pp; English.
CC
CC The present sequence is a cDNA clone encoding wheat sulphate permease, a
CC sulphate assimilation protein. This sequence is obtained from wki
CC library, clone wki.pk0028.e1, derived from wheat seedlings, 1 hour after
CC treatment with fungicide. This sequence is used as a probe to isolate
CC other plant sulphate assimilation proteins, for genetic and physical
CC mapping of related genes and as markers of traits linked to the gene.
CC This is useful for plant breeding and to construct chimeric genes, used
CC to create transgenic plants with altered levels of sulphate permease. The
CC sulphate permease peptides are useful for producing antibodies, that are
CC used to screen and isolate cDNA clones
XX
SQ Sequence 2311 BP; 670 A; 493 C; 537 G; 611 T; 0 U; 0 Other;

Query Match 26.1%; Score 640.4; DB 3; Length 2311;
Best Local Similarity 59.4%; Pred. No. 1.9e-147;
Matches 1104; Conservative 0; Mismatches 751; Indels 3; Gaps 1;
Qy 122 GTGCCAACAGTCGAGGTTCCACGCCACACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTG 181
Db 187 GTGTACAAAGTGGCTATCCCCCTCGAAGAACTTGGCCACAGAGTTACAGAAACATTG 246
Qy 182 AAGGAGACTTTCTTCCCTGATGACCCCTTTGAGGAGCTTCAAGAAACAGCCAGCTTCCAAG 241
Db 247 AGGAGAGCTTTCTTCCACCAACCCGCTGCGTCACTATAAGGGCCAATCCGACCCGAG 306
Qy 242 AAGTTTCATGCTTGGCTTTCAGTCTCTTCTTCCCATTTTTCGAATGGGCTCCCAATACACC 301
Db 307 AGGTTTCATGATGGGCTGGAGTTCTTGTTCCTATATTTGGGTGGGTTAGGATTACAGT 366
Qy 302 TTTTCAGTTCTTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCTCT 361
Db 367 CTCACAAAGTTCAAGGCGATCTGATGCGGAGTGGACCATCGCAAGTCTCTGTATTCTCT 426
Qy 362 CAGGGCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATTTCTGGACTATATTCGAGC 421
Db 427 CAGGACATTGGCTATTTCGAAGCTTGCTAATCTGGATCCGACGATATGGGCTTTACTCCAG 486
Qy 422 TTTATACCACTTGTATTTATCGGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACT 481
Db 487 TTTTCATTTCTCCATTTGATCTATCTGCAATGGGTAGCTCAAGGATATAGCGATTGGTCCA 546
Qy 482 GTGGCGTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCCTTGCATCCCAATGAA 541
Db 547 GTTGTGTGGTTTCTTTTGTAGTTCACTTCTACAAGCTGAGGTTGACCATGTCAAA 606
Qy 542 GACCCAAAGCTTTTACCTCCACCTGGCTTTTACAGCTTACATTTATTTGCTGGTGTTCAG 601
Db 607 AACAAAGGAGGAATACATGGCCTCGCTTTTTCACGGCAACCTTCTTCTGCTGTATCACTCA 666
Qy 602 GCTGCTTGGGTCTGTTTAGTTAGTTGGGTTGATCGTGGATTTTCTGTCACTATCAACCAT 661

Dd	667	GCAGCCTTAGGATTTCTAAGTTAGGATTCCTTATAGAGTTCTTGTGCGATGTCGGATT	726
Qy	662	ATAGGGTTCAATGAGGAGCAGCCCGTGGTGTGCTGTCGAGCAACTAAATCGATCTTT	721
Dd	727	GTCCGAATCATGGGGAGCTGCCATTAATAATTGCCCTGAGCAGCTGAAATACGTGTG	786
Qy	722	GGCCTTTGAGCATTTCAACCCATGAGCTGATATCATATCAGTGTGCGCTCTGTTTTCAAC	781
Dd	787	GGCATCGCAAACTTTACAAGGAAACCGACATAGTTTCTGTATGAAATCTGTCTGGAG	846
Qy	782	CAAACTCAT---GAGTGGAGGTGGGAAAGTGTGTTAGTAGTGTCTTTCAATTTCTTC	838
Dd	847	TCAGTTTCATCAGCGGTGGAACTGCGACAGCAAAATTGTGATTTGGCGTATCTTTTCTCGTGGTTTTC	906
Qy	839	CTCCTTTAGCACAAGATACCTTCAGCAAAACGACCAAGGTTTTTTTGGGTGTCAGCAATG	898
Dd	907	CTTCTGTTTGGAGTACATCGGAAGAGAAAGGAAGCTTTTCTGGGTGCCAGCTATT	966
Qy	899	GGCCCAATTGACGTCCTGTTATTTGGGAAGTCTCTTGGTTTATTTCACTCACGCCGAGAAG	958
Dd	967	GCTCCTATAATTTCACTGATTTCTAGCAACATTTTGTGTATACATTACTGTCGCGACAAG	1026
Qy	959	CACGGTGTGAAGTATAGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAAT	1018
Dd	1027	CAAGGAGTTCAGATAGTGAAGCAATTGAACAGGGAATCAACCCATCATCAGTACACAAG	1086
Qy	1019	CTGGTATTTGTGTCGCTTACATGACTACAGCTGTCAAACTGGCAATTCGTTGGGCATC	1078
Dd	1087	ATTATTTTACCGGCCATTTGTGCAAAAGGTTTCAAGATCGGTGTTGTTCGGGCATA	1146
Qy	1079	ATATCACTTTGCGGAAGGAATAGCAGTAGGAAGGCTTTGCAATGTTATAAAAATTACAA	1138
Dd	1147	GTTGGTTTGACAGAAGCTGTAGCTATTGGAAGGACATTTGCTGCTATGAGGACTACCAG	1206
Qy	1139	ATTGATGGCAACAAAGAGATGATAGCTATTGGGACCAATGAACGTAGTGGTCTTTTCAC	1198
Dd	1207	TTAGATGGAACAAGGAGATGGTAGCACTTTGGAACCAATGAACATAGTGGCTCAATGCA	1266
Qy	1199	TCTTGCTACTCACAACAGACCATTTTCGCGTTGCGCTGTGAACATTAACGCTGGATGC	1258
Dd	1267	TCTTGCTATGTCAACAGGTTCTTTCTCACGTTTCGGCAGTTAACTTCATGCGTGGCTGC	1326
Qy	1259	AAGACAGCAGCTTCCAACTTATAATGTCACTTTGCAGTAATGTTGACATTTGTTATTCCTG	1318
Dd	1327	AAGACTCTGTTATCCAAATGTGGTTATGTAGTAGTGTTCTTCTTACCTTGTGTGTCATC	1386
Qy	1319	ACACCTTGTTCATTAACACTCCCTGGTGGTGTCTATCAGCTATTATGCTATCTGCAATG	1378
Dd	1387	ACACCGCTATTCAAATATACCCGAATGCAATCTTAGGGTCGATCATATTATTTCTGCGGTG	1446
Qy	1379	CTTGGACTCATAGATTATGAAGCAGCCATCCATCTATTAAAGTTGACAAAATTGCACTTT	1438
Dd	1447	ATCGGCTTGTGACTACGAAGCAGCAATTCATCTGGAAGTTGACAAATTTGCACTTTC	1506
Qy	1439	GTGGTGTGATGATGATACATATGGCGTGGTCTTTGGCAGTGTGTGAAATTTGGCTTAGTC	1498
Dd	1507	ATTGCTTGATGGAGCTTTTTCGGTGTGTTTGTGATTCGTTGAGATTGGCCCTCTG	1566
Qy	1499	ATAGCTATTGTAATATCTGACTTCGGGTACTTCTATTATTATGCAAGGCCAGGACATTC	1558
Dd	1567	ATTGCTGTAGCAATCTCATTTGCAAAATACTTCTTCAAGTAACAGGCAAGGACAGCC	1626
Qy	1559	GTTTTGGCAACATTCCAAATTCGTGATATACCGAAATGTTGAGCACTATCAAAATGCA	1618
Dd	1627	CTACTTTGGAACCTTCCCGCACCATATATACCGGAACATACGCCAGTATCCAGMAGCA	1686
Qy	1619	AAACATGTTTCTGGAAATGCTAAATTTCTAGAGATTGATGCACCAATTTTACTTTGCCAATGCC	1678
Dd	1687	AAACTTACTCTGGGGTGTGATTGTGAGGGTTGATTCTGCTATTATTATTTTCCAACTCT	1746
Qy	1679	AGCTATTTAAGAGAAGGATCAACAGGTGATGATGAAGAGAAGAAAGATTTAAAGCT	1738
Dd	1747	AATTACGTCGGAAGAAGAATTTCTTAGTGGCTGACAGACGAAAGAACAGACTAAAAGCA	1806
Qy	1739	ACAGGGGAGACTAGTTTGCAGTATGTTATAATTGATATGAGTCTGTTTGGAAACATTGAT	1798
Dd	1807	GTGGGATTCCCTAAATCAATGTTTCTTGATTTGGAAATGTCGCCGTCACTGCATCTGAT	1866
Qy	1799	ACAAGTGGAAATAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAG	1858
Dd	1867	ACAAGCGGCATACATGCTCTTGAAGATCTATACAAGAATCTTTCAGAAAAAAGATATGAG	1926
Qy	1859	CTTGTGTTTGGTCAATCTCTGTAAGTGAAGTGAAGAAAACCTGAACAATTCGAAGTTCCTCA	1918
Dd	1927	CTCATTTCTGTCAAAATCCTGGTTCCGTGCTATAGAAAAAATCTCAAGCGCTCGAAGCTCAC	1986
Qy	1919	AATCAATTTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGCGCTTGGAGCATGCA	1976
Dd	1987	GAGCATTGGGAAGCAGCAATATATATCTTCGCGGTCTCTGACGCTGTGCGATTCTGTA	2044
RESULT 15			
ABQ82719			
ID	ABQ82719	standard; cDNA; 1971 BP.	
XX	ABQ82719;		
AC	ABQ82719;		
DT	03-JAN-2003	(first entry)	
XX	Arabidopsis thaliana sulphate transporter Sultr1;3 encoding cDNA SEQ:1.		
DE	Arabidopsis thaliana; sulphate transporter; chromosome 1; plant;		
XX	sulphur-containing metabolite; glutathione; phytokeatin; gene; ss.		
KW	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
XX	Key	Location/Qualifiers	
FT	CDS	1..1971	
FT		/tag= a	
FT		/product= "sulphate transporter Sultr1;3"	
XX	JP2002272472-A.		
PN			
XX	24-SEP-2002.		
PD			
XX	22-MAR-2001; 2001JP-00082891.		
PF			
XX	22-MAR-2001; 2001JP-00082891.		
PR			
XX	(KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.		
PA			
XX	WPI; 2002-718707/78.		
DR	P-PSDB; ABP53777.		
DR			
XX	A gene encoding sulfate ion transporter Sultr1;3, a polynucleotide, an		
PT	oligonucleotide, a recombinant vector, a plant.		
FT			
XX	Claim 2; Page 6-9; 12pp; Japanese.		
PS			
XX	The present invention describes a gene which is present in the first		
CC	chromosome of Arabidopsis thaliana and encodes sulphate ion transporter		
CC	Sultr1;3 having the amino acid sequence given in ABP53777. Also		
CC	described: a polynucleotide purified from the genomic DNA, the mRNA, the		
CC	cDNA or their complementary sequences; an oligonucleotide hybridising		
CC	with the above gene or the above polynucleotide and consisting of a base		
CC	sequence of at least 10 base pairs; a recombinant vector carrying the		
CC	above polynucleotide; and a plant in which the above gene is introduced		
CC	or its descendant or their tissues. The plant can be used for the		
CC	preparation of sulphur-containing metabolites such as glutathione and		
CC	phytokeatin. The present sequence encodes Arabidopsis thaliana sulphur		
CC	transporter Sultr1;3, from the present invention		
XX	Sequence 1971 BP; 504 A; 465 C; 427 G; 575 T; 0 U; 0 Other;		
SQ			
Query Match 25.9%; Score 634.4; DB 6; Length 1971;			
Best Local Similarity 59.1%; Pred. No. 5.5e-146;			

Matches 1104; Conservative 0; Mismatches 761; Indels 3; Gaps 1;			
Qy	122	GTGCACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAGAGTCTCTAAAGTACTCTTTG	181
Db	88	GTCCACAAAGTCGAAGTTCCTCCTTAAGCAAAACCTTTTCAATGAGTTTCATGACACTTTT	147
Qy	182	AAGGAGACTTCTTCCCTGATGACCCCTTGGGAGGTTTCAAGAACAGCCAGCTTCCAAG	241
Db	148	AAAGAAACTTCTTCCACGATGATCCTCTAAGGCATTTCAAGGACCAAGTCAAAATCCAAA	207
Qy	242	AAGTTCATGCTTGCCCTTCAAGTCTTCTCTCCCAATTTTCAATGGGCTCCCAATACACC	301
Db	208	AAGCTCATGCTCGGTATCCAGTCGCTCTTTCGGTATCGAGTGGGGAAGAAATATAAT	267
Qy	302	TTTCAGTCTTGAAGCTGACCTCATAGCTGGGCATCAACATCGCTAGCTTGGCCATTCCT	361
Db	268	CTTAAAGTTGTTTCGCGCGCATCTTATTTGCGGGTTTAAACCATAGCAGTCTCTGCAATCCT	327
Qy	362	CAGGSCATGAGTTATGCAAGCTCGCCAACTCCCTCCCAATTTCTTGGACTATATTCGAGC	421
Db	328	CAGGATATGGATATGCAAGCTCGCGAGTCTTGACCCCTAAGTATGAGTCTATATTCAGT	387
Qy	422	TTTATACCAACATTTGATTTATGGCATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACT	481
Db	388	TTTGTCTCTCGTTGGTGTACGCATGTATGGGAAGCTCAAGGATATAGCGATGGACCC	447
Qy	482	GTGCGGTTGGATCGCTTCTGTAGTGGGTTTCGATGTGAGTAAATGCCGTGTATCCCAATGAA	541
Db	448	GTTCAGTGGTTTCACTCTTATAGTACTCTGCTCGTGTGATCGACCGACCCCAACACA	507
Qy	542	GACCCAAAGCTTTACCTCCACCTGGCTTTCACAGCTACATTTATGCTGGTGTCTTTTCAG	601
Db	508	AACCCATGAATATCTCCCGCTTAGCCCTTCAACCTCCACGTTCTTTGCCGGTGTCACTCAA	567
Qy	602	GCTGCCCTTGGCTCTGTTTAGTGGGTTGATCGTGATTTTCTGCATGCAACCAATA	661
Db	568	GCAGCACTCGATTTCTCAATTTGGGTTTCTTGATCGATTTCTTGCCACGGCGGGTG	627
Qy	662	ATAGGGTTTATGGAGAGCAGCACCGTGGTGTCTCGAGCAACTAAATCGATTCCT	721
Db	628	GTAGGTTTCAATGGCGGAGCAGCCATCACCATTTGCGCTGCAACAGCTCAAAGGCTTCCTA	687
Qy	722	GGCTTGAGCATTTCAACCATGGAGCTGATATCATATCAGTGATCGGCTCTGTTTTCACC	781
Db	688	GGATCAATAAGTTTCAAAAGAAACCGATATCATCGGTTCTTTCTCCGTAATCAGC	747
Qy	782	CAAAAC---TCATGAGTGGAGTGGGAAAGTGTGTGTAGAGTGTGCTTTCATTTTCTTC	838
Db	748	TCAGCCCATCACGATGGAAATGGCAGACAATCTCATTTAGTGATCGTTCTTGATCTTC	807
Qy	839	CTCCTTAGCACAAGATACTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATG	898
Db	808	CTTCTCATCTCCAAGTTTATCGGGAGAGAAAACAAGAAACTGTTTTGGATTCCAGCTATT	867
Qy	899	GCGCCATTGACGTCCTGTTATTTGGGAAGTCTCTTGGTTTATTTTCACTCAGCGCGAGAAG	958
Db	868	GCTCCGGTTAGTATCTGTCTATCATTTCAACCTTCTTGTCTACATAACCCGAGCGCGACAAG	927
Qy	959	CACGGTGTCAAGTGTAGGAGAACTCAAGAGGGTTTGAATCCACCATCACTCACAAAT	1018
Db	928	AAAGGAGTTCAATAGTGAACAATCTTGACAAGAGTCTGAACCTTCTCTTTGGGTCTA	987
Qy	1019	CTGGTATTTGTGTGCGCTTACATGACTACAGCTGTCAAAACTGGCAATGTGTTGGCATC	1078
Db	988	ATATATTTCTCGGCGGATTTACCTTCTCAAGGGCTTCGATAGGCGTTGTCTCAGGCAATG	1047
Qy	1079	ATATCACTTCGGNAGGAATAGCAAGTAGGAAGAGCTTTTGCATGTATAAAATTAACAAT	1138
Db	1048	GTTGCTCTTGAACGAGCTGTAGCGATAGGAAGAACTTTTGCAGCAATGAAAGACTACCAA	1107
Qy	1139	ATTGATGGCAACAAGAGATGATAGTATTTGGGACCACTGAACGTAGTTGGTCTTTTCACC	1198
Db	1108	ATCGATGGTAACAAGAGATGGTAGCAATTAGGCAATGAACGTAATCGGTTCAATGACC	1167

Search completed: August 31, 2005, 03:36:07

Job time : 1320 secs

Qy	1199	TCTTGTCTACCTCAACACAGGACCATTTTTCGGTTCGGCTGTGAACATATAAACGTGTGATGC	1258
Db	1168	TCTTGTCTATGTATCCACCGGTTCTTCTCAAGATCCGCCGCTCAACTTTTATGGCGGATGT	1227
Qy	1259	AAGACAGCAGCTTCCACCAATTAATAGTCTACTTTGCGAGTAAATGTTGACATTTGTTTCCCTG	1318
Db	1228	CAAAAGGCAAGTCTCCAAACATCATGTCCATTTGTCGTCCTCTTAAAGCGTCTCTCTTCCCTC	1287
Qy	1319	ACACCTTGTGTCCATTTACACTCCCTCGGTGGTGTCTATCAGCTATTTATFCGTATCTGCAATG	1378
Db	1288	ACTCTCTCTTTTCAAAATACACACCACCAAGCAATTTCTTGCAGCGATCATCATCAAGCGTGTG	1347
Qy	1379	CTTGGAGCTCATAGATTTATGAAGCAGGCATCCATCTATTTTAAAGGTTGACAAATTTGACTTT	1438
Db	1348	ATTCTCTTGGTTGACGCTTAATGTACCATTTTGTATCTTCAAGATCGATAAGCTCGATTTT	1407
Qy	1439	GTGCTGTGCATGATGATCATACATTTGGCGTGGTCTTTTGGCAGTGTGAAATTTGGCTTAGTC	1498
Db	1408	GTTGCTGTATAGGGGCTTTTGTGTGTCATCTTTGTATCCGTAGAGATTTGGCTTCTA	1467
Qy	1499	ATAGCTATTTGTAATATCTGTACTTTCGGGTACTTCTATTTTATTCAAAGGCCAAGGACATTC	1558
Db	1468	ATAGCGTGGGATATCTTTTGCAGAGATCTTTTGCAGGTTACGAGACCTAGGACAGCG	1527
Qy	1559	GTTTTGGGCAACATTTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCA	1618
Db	1528	ATTCTTTGGAAGATACAGGGACTTCGGTTTACAGGAATATCAATCAGTATCTCTGAAGCG	1587
Qy	1619	AAACATGTTCTCGAATGCTAATTTCTAGAGATTTGATGCACCAATTTTACTTTTGCCAAATGCC	1678
Db	1588	ACTAGGATTTCCGGAGTTTGTGCAATTCGTGTTGATTCGCGGATTTACTTTCTCAAATCC	1647
Qy	1679	AGCTATTTAAGAGAAAGGATCAACAGTGGATTGATCAAGAAAGAAAGAAATTTAAAGCT	1738
Db	1648	AATTATGTTAGGAAAGGATTCAGAGATGGTTGACAGATGAAGAGAGATGGTGGAGCT	1707
Qy	1739	ACAGGGAGACTAGTTTTCAGATGATTTAATTTGATATGAGTGTCTGTTGGAAACATTTGAT	1798
Db	1708	GCAAGATTTGCTAGGATCCAGTTTCTTATCATCGAAATGTCACTGTGTACGGACATCGAT	1767
Qy	1799	ACAAGTGGAAATAGTATGCTTGAAGAGGTGCAAGAAGATTACAGAGAGAGAGACTACAG	1858
Db	1768	ACTAGTGGTATTCACGCTTTAGAAAGCTTTGTATAGTCTCTCCAAAACAGAGACATTCAG	1827
Qy	1859	CTTGTGTTTGGTCAATCTCTGTAAGTGAATGATGAAGAACTGAACAAATCGAAGTTCCAA	1918
Db	1828	TTGGTTCTAGCGAATCCAGGACCGCGGTCATAAAATAGCTACATGTTTCTCACTTTGCG	1887
Qy	1919	AATCATTTTGGGAAGAAATGGATCTATCTGACTGTGTTGAAGAGCGCGTTGGAGGATGCAAC	1978
Db	1888	GACTTGATAGGACACGACAAAATCTTTTCTGACGGTGGCTGAGGCCGTTGATTTCTTGCTCC	1947
Qy	1979	TTCAATCT 1986	
Db	1948	CTTAAACT 1955	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 02:38:26 ; Search time 429 Seconds
(without alignments)
9340.890 Million cell updates/sec

Title: US-10-762-049-17
Perfect score: 2449
Sequence: 1 gcacagctagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata1/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata1/ina/5B_COMB.seq.*
- 3: /cgn2_6/prodata1/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata1/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata1/ina/PCUTUS_COMB.seq.*
- 6: /cgn2_6/prodata1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2449	100.0	2449	4	US-09-720-317A-17
2	765.2	31.2	1981	4	US-09-720-317A-3
3	720.6	29.4	2067	4	US-09-720-317A-15
4	655.4	26.8	2279	4	US-09-720-317A-1
5	640.4	26.1	2311	4	US-09-720-317A-19
6	324	13.2	780	4	US-09-720-317A-7
7	216.2	8.8	1240	4	US-09-720-317A-5
8	205.6	8.4	2022	4	US-09-720-317A-21
9	117.4	4.8	493	4	US-09-720-317A-13
10	117.4	4.8	484	4	US-09-720-317A-9
11	90.4	3.7	2441	4	US-09-785-381-4
12	82.6	3.4	4927	4	US-09-949-016-5627
13	79.8	3.3	4113	4	US-09-785-381-2
14	77.6	3.2	2487	4	US-09-248-796A-6643
15	76.4	3.1	7218	1	US-08-232-463-14
16	64.4	2.6	2682	4	US-09-614-221A-222
17	63.8	2.6	2913	4	US-09-795-927-6
18	63.8	2.6	3749	4	US-09-795-927-8
19	58.2	2.4	510	4	US-09-720-317A-11
20	58	2.4	758	4	US-09-270-767-470
21	58	2.4	758	4	US-09-270-767-15752
22	57.8	2.4	3061	1	US-09-570-842-1
23	57.8	2.4	3061	1	US-08-700-576-1
24	56	2.3	1773	4	US-09-902-540-7746
25	56	2.3	4854	4	US-09-902-540-768
26	53.4	2.2	2832	4	US-09-949-016-4516
27	53.4	2.2	8774	4	US-09-949-016-16258

ALIGNMENTS

RESULT 1

US-09-720-317A-17

; Sequence 17, Application US/09720317A

; Patent No. 6696292

; GENERAL INFORMATION:

; APPLICANT: Stephen M. Allen

; APPLICANT: Saverio C. Falco

; APPLICANT: Catherine J. Thorpe

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167

; CURRENT APPLICATION NUMBER: US/09/720,317A

; CURRENT FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 14-07-1998

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 17

; LENGTH: 2449

; TYPE: DNA

; ORGANISM: Glycine max

US-09-720-317A-17

Query Match 100.0%; Score 2449; DB 4; Length 2449;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGCTAGCTCGCACATTAAGTTATATACACATATTTGCTTCTTAGAAATACTAT 60

Db 1 GCACGAGCTAGCTCGCACATTAAGTTATATACACATATTTGCTTCTTAGAAATACTAT 60

QY 61 TATTGAAGATATGGGAGTGATAGTTATGAGTACCCCTTTGGGCATCAACACTTTCAGAG 120

Db 61 TATTGAAGATATGGGAGTGATAGTTATGAGTACCCCTTTGGGCATCAACACTTTCAGAG 120

QY 121 AGTGACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAGAGTCTCTAAAGTACTCTT 180

Db 121 AGTGACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAGAGTCTCTAAAGTACTCTT 180

QY 181 GAAGGAGACTTCTTCCCTGTAGTACCCCTTTGAGGCGAGTTCAAGAACAGCAGCTTCCAA 240

Db 181 GAAGGAGACTTCTTCCCTGTAGTACCCCTTTGAGGCGAGTTCAAGAACAGCAGCTTCCAA 240

QY 241 GAAGTTCATGCTTGGCCTTCAGTCTTCTTCCCATTTTCGAATGGCTCCCAATATACAC 300

Db 241 GAAGTTCATGCTTGGCCTTCAGTCTTCTTCCCATTTTCGAATGGCTCCCAATATACAC 300

QY 301 CTTTCAGTCTTGAAGAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCC 360

Db 301 CTTTCAGTCTTGAAGAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCC 360

QY	361	TCAGGCGATCAGTTATGCCAAGCTCGCCAACTCCCTCCCAATCTTTGGACTATATTCGAG	420
DB	361	TCAGGCGATCAGTTATGCCAAGCTCGCCAACTCCCTCCCAATCTTTGGACTATATTCGAG	420
QY	421	CTTTATACCAACCATTAATTTATCGGATGATGGGTAGCTCGAGGATTTGGCAGTGGGAC	480
DB	421	CTTTATACCAACCATTAATTTATCGGATGATGGGTAGCTCGAGGATTTGGCAGTGGGAC	480
QY	481	TGTGGCGTTGGATCGCTCTGATGSGTTCGATGTTGAGTAATGCGGTTGATGCCCAATGA	540
DB	481	TGTGGCGTTGGATCGCTCTGATGSGTTCGATGTTGAGTAATGCGGTTGATGCCCAATGA	540
QY	541	AGACCCAAAGCTTTACCTCCACCTGGCTTTTCACAGCTACATTAATTTGCTGGTGTTC	600
DB	541	AGACCCAAAGCTTTACCTCCACCTGGCTTTTCACAGCTACATTAATTTGCTGGTGTTC	600
QY	601	GGCTGCTTTGGGTCTGTTTAGGTTGGGTTGATCGTGGATTTCTGTCCATGCCAAT	660
DB	601	GGCTGCTTTGGGTCTGTTTAGGTTGGGTTGATCGTGGATTTCTGTCCATGCCAAT	660
QY	661	AATAGGGTTTCATGGGAGGAGCAGCCAGGTGGTGTCTGTCAGCACTAAATCGAATCT	720
DB	661	AATAGGGTTTCATGGGAGGAGCAGCCAGGTGGTGTCTGTCAGCACTAAATCGAATCT	720
QY	721	TGSCCTTGAGCAATTCACCCATGGAGCTGATATCATATCAGTGAAGCGCTCTGTTTCAC	780
DB	721	TGSCCTTGAGCAATTCACCCATGGAGCTGATATCATATCAGTGAAGCGCTCTGTTTCAC	780
QY	781	CCNAATCATGATGAGGTGGGAAAGTGTGTGTTAGGATGCTTCATTTCTTCT	840
DB	781	CCNAATCATGATGAGGTGGGAAAGTGTGTGTTAGGATGCTTCATTTCTTCT	840
QY	841	CCTTAGCACAAGATCTTCAGCAAAACCGACCAGGTTTTTTGGGTGTCAGCAATGCG	900
DB	841	CCTTAGCACAAGATCTTCAGCAAAACCGACCAGGTTTTTTGGGTGTCAGCAATGCG	900
QY	901	GCCATTGACGTCGGTTATATTTGGGAAGTCTCTTTGGTTATTTCACTCAGCCGAGAGCA	960
DB	901	GCCATTGACGTCGGTTATATTTGGGAAGTCTCTTTGGTTATTTCACTCAGCCGAGAGCA	960
QY	961	CGGTGTTGAAGTATAGGAACCTGAAGAGGGTTTGAATCCACCATCTACAAATCT	1020
DB	961	CGGTGTTGAAGTATAGGAACCTGAAGAGGGTTTGAATCCACCATCTACAAATCT	1020
QY	1021	GGTATTTGTCGCCCTTACATGACTACAGCTGTCAAACTGGCATTTGCTGGCATCAT	1080
DB	1021	GGTATTTGTCGCCCTTACATGACTACAGCTGTCAAACTGGCATTTGCTGGCATCAT	1080
QY	1081	ATCATTGCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATPAAAAATTACAATAT	1140
DB	1081	ATCATTGCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATPAAAAATTACAATAT	1140
QY	1141	TGATGCGAACNAAGATGATAGCTATTTGGGACCATGAACGTAGTTGGTTCTTTCACCTC	1200
DB	1141	TGATGCGAACNAAGATGATAGCTATTTGGGACCATGAACGTAGTTGGTTCTTTCACCTC	1200
QY	1201	TTGCTACCTCACAAAGAGCAATTTTCGGCTTCGGCTGTGAACATAAAGCTGGATGCAA	1260
DB	1201	TTGCTACCTCACAAAGAGCAATTTTCGGCTTCGGCTGTGAACATAAAGCTGGATGCAA	1260
QY	1261	GACAGAGCTTCCAAATTAATATGATCTGAGTAATGTTGACATTTGTTATCTCTGAC	1320
DB	1261	GACAGAGCTTCCAAATTAATATGATCTGAGTAATGTTGACATTTGTTATCTCTGAC	1320
QY	1321	ACCTTTGTTCCATTACACTCCCTGGTGGTGTATCATGCTATATCGTATCTGCAATGCT	1380
DB	1321	ACCTTTGTTCCATTACACTCCCTGGTGGTGTATCATGCTATATCGTATCTGCAATGCT	1380
QY	1381	TGGACTCATAGATTATGAAGAGCCCATCATCTATTTAAAGTTGACAAATTTGACTTTGT	1440
DB	1381	TGGACTCATAGATTATGAAGAGCCCATCATCTATTTAAAGTTGACAAATTTGACTTTGT	1440
QY	1441	GGTGTGCATGAGTGCATACATTTGGGCGTGTCTTTTGGCAGTGTGTGAATTTGGCTTAGT	1500

DB	1441	GGTGTGCATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGT	1500
QY	1501	AGCTATTGTAATATCTGTACTTTGGGTACTTCTATTATTGCAAGCCCAAGACATTCT	1560
DB	1501	AGCTATTGTAATATCTGTACTTTGGGTACTTCTATTATTGCAAGCCCAAGACATTCT	1560
QY	1561	TTTGGGCAACATTCGAAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATGCAAA	1620
DB	1561	TTTGGGCAACATTCGAAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATGCAAA	1620
QY	1621	ACATGTTCTGGAATCTTAATTTCTAGAGATTGATGCCAAATTTACTTTGCCAATGCCAG	1680
DB	1621	ACATGTTCTGGAATCTTAATTTCTAGAGATTGATGCCAAATTTACTTTGCCAATGCCAG	1680
QY	1681	CTATTTAAGAGAAAGATCACAAGGTGATGATGAAGAGAAAGAAAGATTAAGACTAC	1740
DB	1681	CTATTTAAGAGAAAGATCACAAGGTGATGATGAAGAGAAAGAAAGATTAAGACTAC	1740
QY	1741	AGGGGAGACTAGTTTGCAGTATGTTATTAATCATATGATGCTGTGGAAACATTTGATAC	1800
DB	1741	AGGGGAGACTAGTTTGCAGTATGTTATTAATCATATGATGCTGTGGAAACATTTGATAC	1800
QY	1801	AAGTGGAAATTAATGCTTTGAAAGAGGTGAAGAAATTAAGAGAGAAAGAGCTACAGCT	1860
DB	1801	AAGTGGAAATTAATGCTTTGAAAGAGGTGAAGAAATTAAGAGAGAAAGAGCTACAGCT	1860
QY	1861	TGTTTTGCTCAATCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAA	1920
DB	1861	TGTTTTGCTCAATCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAA	1920
QY	1921	TCATTTAGGGAAGAAATGGATCTATCTGACTGTTCAAGAGGCGTTGGAGCATGCAACT	1980
DB	1921	TCATTTAGGGAAGAAATGGATCTATCTGACTGTTCAAGAGGCGTTGGAGCATGCAACT	1980
QY	1981	CAATCTAGCTGCAAGCAAAACGAAACCAAGAAAGATGAAACAGAGGTTGAAACAAAT	2040
DB	1981	CAATCTAGCTGCAAGCAAAACGAAACCAAGAAAGATGAAACAGAGGTTGAAACAAAT	2040
QY	2041	GTGACTGATCATATGCCAAGAGTATTTCTAATAACTCAAAAGCTTATTCGTTTCT	2100
DB	2041	GTGACTGATCATATGCCAAGAGTATTTCTAATAACTCAAAAGCTTATTCGTTTCT	2100
QY	2101	CTTTAGTATTTTACCACTACAAATGTGTCATGAGAAATTTCTGAATCAGCCGAAAGAGT	2160
DB	2101	CTTTAGTATTTTACCACTACAAATGTGTCATGAGAAATTTCTGAATCAGCCGAAAGAGT	2160
QY	2161	TTTAAAGGCATAGGAAAAATGAAAGATGCAAGGCTTCTTAATTTCTCAACTCTGCATCT	2220
DB	2161	TTTAAAGGCATAGGAAAAATGAAAGATGCAAGGCTTCTTAATTTCTCAACTCTGCATCT	2220
QY	2221	TAGTTAGAGAAAAATCTCCTATGATAGGCTGTTGAAATAATCTTTACGTATCATGCTTG	2280
DB	2221	TAGTTAGAGAAAAATCTCCTATGATAGGCTGTTGAAATAATCTTTACGTATCATGCTTG	2280
QY	2281	ATAATATATTTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACACATG	2340
DB	2281	ATAATATATTTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACACATG	2340
QY	2341	TAAAGAGTAAAGAGTGTGTTGCTAGCACTCCTCCATATTCATTTGAAAGTAAATGGC	2400
DB	2341	TAAAGAGTAAAGAGTGTGTTGCTAGCACTCCTCCATATTCATTTGAAAGTAAATGGC	2400
QY	2401	ATGAGAAATTTAAATAATCTTTTGGAAAAAATTTTGGAAAAAATTTTGGAAAAA	2449
DB	2401	ATGAGAAATTTAAATAATCTTTTGGAAAAAATTTTGGAAAAAATTTTGGAAAAA	2449

RESULT 2

US-09-720-317A-3

; Sequence 3, Application US/09720317A

; Patent No. 6696292

; GENERAL INFORMATION:

APPLICANT: Stephen M. Allen
APPLICANT: Saverio C. Falco
APPLICANT: Catherine J. Thorpe
TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
FILE REFERENCE: BB-1167
CURRENT APPLICATION NUMBER: US/09/720,317A
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/092,833
PRIOR FILING DATE: 14-07-1998
NUMBER OF SEQ ID NOS: 31
SOFTWARE: Microsoft Office 97
SEQ ID NO 3
LENGTH: 1981
TYPE: DNA
ORGANISM: Zea mays
US-09-720-317A-3

Query Match 31.2%; Score 765.2; DB 4; Length 1981;
Best Local Similarity 65.8%; Pred. No. 4.3e-202;
Matches 1112; Conservative 0; Mismatches 578; Indels 0; Gaps 0;
313 GAAAGCTGACCTCATAGCTGGCATACCATCGCTAGCTTGCCCATTCCTCAGGCGATCAG 372
Db 7 GGAGTCCGACTGATCGCGGCATCACCATCGCCAGCCTCGCCATCCCGCAGGSCATCAG 66
373 TTATGCCAAGCTGCCAAACCTCCCTCCAATTTCTGGACTATATTCGAGCTTTATACACC 432
Db 67 CTAGGCCAAGCTGCCAAACCTCGCGCCCGTGTCTGGACTCTACTCGAGCTTCGTGCGGCC 126
433 ATTGATTATATGCGATGATGGTAGCTCGAGGGATTTTGGCAGTGGGACTGTGGCGGTGG 492
Db 127 GCTGGTGTAGCGCTGATGGGAGCTCCAGGACCTCGCGGTGGGAGCGGTGGCGGTGGC 186
493 ATCGCTTCTGATGGGTTGATGTTGAGTAATGCGGTTGATCCCAATGAAGACCCAAAGCT 552
Db 187 GTCGTGCTCATCAGCTCCATGCTCGSCAGGAGTGTGCCGACGAGAACCCCGTGTCT 246
553 TTACTCCACCTGGCTTTCACAGTACATATTTTGTGCTGGTGTGTTTTCAGGCTGCTGGG 612
Db 247 CTACTGTCACCTGCGCTTTCACCGCACCTTCTTTCGCGCGGCTCTTCAGGCGCTCGCTCGG 306
613 TCTGTTTATGTTGGGTTGATCGTGGATTTTCTGTACATGCAACCATATAGGTTTCAT 672
Db 307 CTTCTCAGGTTGGGTTTCATGTTGGACCTGCTGTGCGCAGCAGATCGTGGGTTTCAT 366
673 GGGAGGAGCAGCAGCGTGGTGTCTGTCGACAACTAAATTCGATTTCTGGCCCTTGAGCA 732
Db 367 GCGCGCGCGCGACGCTGGTGTGCTGCTGACAGCTGAAGGCGATGCTGGGCTCGTCCA 426
733 TTTACCCATGGAAGCTGATATCATATCAGTGATGCGCTCTGTTTTCACCAAACTCATGA 792
Db 427 CTTTACCCACCTCCACCGACGCTGCTCTCGTCAATGGAATCGGTTCTTACGCCAGACACCA 486
793 GTGAGTGGGAAAGTCTGTTTGTAGATGCTGTTTCATTTCTCTCTCTCTAGCACAAG 852
Db 487 GTGCGGTGGGAGGCTCTGCTGCTGCGCTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCT 546
853 ATACTTCAGCAAAAACGACCAAGGTTTGTGGGTGTCAGCAATGGCGCATTTGACGTC 912
Db 547 CTTTATCAGCAAGAGGCTCCCAAGCTGTTCTGATCTCTCGCGCGCGCGCTGTGAGCTC 606
913 CGTTATATGGGAAGTCTCTTGGTTTATTTTCTACGCGCGAGAGCAAGCGGTGTTGAAGT 972
Db 607 CGTCTGCTCGGAGCGCTTCTGCTGTGTACCTCAGCGACGCTGAAACCCACGCGCATCGAAGT 666
973 GATAGGAGAACTGAGAGGTTTGAATCCACCATCACTCACAAATCTGGTATTTGTGTC 1032
Db 667 GATCGGTTACTGAGAAAGGCTGAAATCCACCGTCTGGTGACAGGCTGCAATTTCTCACC 726
1033 GCCTTACATGACTACAGCTGTCAAACTGGCATTTGCTGGCATCATATCACTTTGCGGA 1092
Db 727 GCCCTACATGATGCTCGCGCTCAAGACTGGGATCATCACCGGCTCATTTGCCCTCGCCGA 786

Qy 1093 AGGAATAGCAGTAGGAGAAGCTTTGCAATGTATATAAAATTAACAATATTGATGGCAACAA 1152
Db 787 AGGAATCGCGTGGGAGGAGCTTCGCCATTTTCAAGAACTACACATGACGGAACAA 846
Qy 1153 AGAGATGATAGCTATTGGGACCATGAACGTAAGTGTGTTCTTTCACCTCTTGTCTACCTCAC 1212
Db 847 GGAGATGATCGCATCGGACGATGAACGTCCTGGGCTCGCTACGCTGCTGCTACCTGAC 906
Qy 1213 AACAGGACCAATTTTCGGCTTCGGCTGTGAACATATAACGCTGGATGCAAGACAGCAGCTTC 1272
Db 907 CACGGGGCCCTTCTCGCGCTCCCGGTGAACACTACAACGCGGGTGCAGGACGCGCATGTC 966
Qy 1273 CAACATTAATGATGCTGCTGCACTTAATGTTGACATTTGTTATTTCTGTGACACCTTGTTC 1332
Db 967 GAAAGTGGTCAATGCTGCTGGCGGTGATGGTCAAGCTGCTTCTGACGCGCTGTTC 1026
Qy 1333 TTACACTCCCTCGTGTGCTATCAGCTATTATCGTATCTGCAATGCTTGGACTCATAGA 1392
Db 1027 CTACACGCGCTGGTGTGCTGTCGGGATCATGCTCTCGCGATGCTGGGCTGTGTCGA 1086
Qy 1393 TTATGAAGCAGCCATCCATCTATTAAAGGTGTGAACAATTTGACTTTGTGTGTGCTATGAG 1452
Db 1087 CTTTCGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146
Qy 1453 TGCATACATTTGGCGTGTCTTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTATTGTAAT 1512
Db 1147 CGGTAACCTGGGCGTGTCTTTCGGCAGCGTCAAGGTGCGGCTGCTGCTGCTGCTGCTGCTG 1206
Qy 1513 ATCTGTACTTTCGGGTACTTCTATTATTTGCAAGGCCAAGGACATTTGTTTTGGGCAACAT 1572
Db 1207 CTCCTGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266
Qy 1573 TCCAAATTTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATGTTCTCTGG 1632
Db 1267 CCCCGCACCATGCTGTACCGGAGGATGGACAGTACGCGCGCGCAGACGCTGCTGCTGCTG 1326
Qy 1633 AATGCTAATTTCTAGATTTGATGCAACAATTTTACTTTGCAATGCGAGCTATTTAAGAGA 1692
Db 1327 CGTGTCTGCTGCGGCTGCGACGCGCGCTGCTTCTTTCGCAACGCCAGCTACCTGCGAGA 1386
Qy 1693 AAGGATCACAAAGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACAGGGAGACTAG 1752
Db 1387 GAGGATCTCGCGTGGATGCAAGCAGGAGGAGCGCAACAGAGCCAGGCGGAGATGGG 1446
Qy 1753 TTTGCAATTTGTTAATTTGATATGATGCTGTTTGGAAACATTTGATACAAAGTGGAAATAG 1812
Db 1447 CGTGGGTAGCTGTTGCTCTCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
Qy 1813 TATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGCTACAGCTTGTGTTGTTGCTCAA 1872
Db 1507 CATGCTGGACGAGCTCAACAAGTCTTTGGACAGGAGGGAATGCAGATCGTGTGCTGCGAA 1566
Qy 1873 TCCGTGTAAGTGAAGTCAATGAAGAACTGAACAAATCGAAGTTCCAAATTCATTTAGGGAA 1932
Db 1567 CCGGCGCAGGAGATCATGAAGAGCTGAGACAGCTCAAGGTGCTGAGGAGAGATCGGCCA 1626
Qy 1933 GAAATGGATCTATCTGACTGTTTGAAGAGCGCTTGGAGCATGCAACTTCAATCTACGTGC 1992
Db 1627 CGAGTGGGTGTTCCCGACGCTGGCGAGCGGTGGCGTCTGTCGCGACTACGTGCTGCACTC 1686
Qy 1993 AGCAAAACG 2002
Db 1687 GCACAAGCCG 1696

RESULT 3
US-09-720-317A-15
; Sequence 15, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins									
; FILE REFERENCE: BB-1167									
; CURRENT APPLICATION NUMBER: US/09/720,317A									
; CURRENT FILING DATE: 2000-12-21									
; PRIOR FILING DATE: 2000-12-21, 833									
; PRIOR FILING DATE: 14-07-1998									
; NUMBER OF SEQ ID NOS: 31									
; SOFTWARE: Microsoft Office 97									
; SEQ ID NO 15									
; LENGTH: 2067									
; TYPE: DNA									
; ORGANISM: Glycine max									
US-09-720-317A-15									
Query Match 29.4%; Score 720.6; DB 4; Length 2067;									
Best Local Similarity 62.2%; Pred. No. 1.1e-189;									
Matches 1134; Conservative 0; Mismatches 689; Indels 0; Gaps 0;									
QY	143	CGGCCACAGCCGTTTTCAAGTCTCTAAAGTACTCTTTGAAGGAGACTTTTCTTCCCTGAT	202						
DB	8	CCACACCAGACACACTCCACAAACTTAGGCACAGAGTCTCCGAATCTTCTCCAGAT	67						
QY	203	GACCCCTTGAGGAGTTCGAAGAACAGCCAGCTTCCAAGAGTTCAATGCTTGGCCCTTCAG	262						
DB	68	GACCCCTCTCCACCGTTTTCAAGAACCAAACTCGGCTTTAAAAAGTTCTCTCGCAGCTTCAG	127						
QY	263	TTCTTCTTCCCAATTTTCAAGTGGGCTCCCAATACACCTTTCAGTTCTTGAAGCTGAC	322						
DB	128	TATCTCTTCCCAATTTTCAAGTGGGCTCCCAATACACCTTTCAGTTCTTGAAGCTGAC	187						
QY	323	CTCATAGCTGGCATCACCTCGTAGCTTGGCCATTCCTCAGGGCATCAGTTATGCCAAG	382						
DB	188	CTCATCTCTGGCTCACCTTGCCAGCTCGCCATTCCTCAGGGAATCAGTTATGCCAAG	247						
QY	383	CTCGCCAACTCCCTCAATCTTGAGCTATATTCGAGCTTTATACCAACATGATTTAT	442						
DB	248	CTTGCCAACTTGCACCTATCTCTGATATATTCGAGTTTGTTCCTCCCATTTGATATAC	307						
QY	443	GGCATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTGTGGCGTTTGGATCGCTTCG	502						
DB	308	TGCTGCTTGGAAAGTTCTAGACATCTTGGTTGGACCTGTTCATTCGGTCTTTGGTC	367						
QY	503	ATGGGTTCGATGTGAGTAATCCGTTGATCCCAATGAAGACCCAAAGCTTTACCTCCAC	562						
DB	368	ATGGGATCAATGTTAAGTGATATAAATTTCTTACACTCAAGAACCTATTCTCTATCTGGGA	427						
QY	563	CTGGCTTTCACAGCTACATTTTGTGTTGTTTTCAGGCTGCCCTGGTCTGTTTAGG	622						
DB	428	TTGGCTTTTACCGCCACTTTCTTTGCTGTTGTTATTCGAAGCTTCTCTGGGTATATTTAAGG	487						
QY	623	TTGGGTTTCATCGTGGATTTTCTGTACATGCACATGCAACCAATAATAGGGTTTCATGGGAGCA	682						
DB	488	CTAGGCTTCGTAAATGATTTTCTGTCGAAGCAACGCTGGTTGGATTCACGGCGTGT	547						
QY	683	GCCACGGTGGTGTCTGCAGCAACTATAATCGATTTCTTGGCTTGAGCAATTTCAACCCAT	742						
DB	548	GCCATTTATTTGTCTACTGTCAGCAGCTGAAAGGTTTACTTTGGAATAGTGCACTTTACCAGC	607						
QY	743	GGAGCTGATATCATATCAGTATGCGCTGTGTTTCAACCAACTCATGAGTGGAGGTGG	802						
DB	608	AAGATGCAATAATATCCAGTAACGATCTCTGTTTTCGAAGCAAGACACAGTGGTCAATGG	667						
QY	803	GAAAGTGTGTGTAGGATGTGCTTCATTTTCTCTCTAGCAACAAGTACTTTCAGC	862						
DB	668	CAAAACCATTTTGGGATTCGGCTTCTGTTCTTCTGCTGACAAACAGGCACATTAGT	727						
QY	863	AAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCCCAATGACGTCGGTTATTTAGT	922						
DB	728	TTGAGGAACCAAAACATTTCTGGGTTTCAGCAGCTGCCCAATGACATCAGTTATTTCTG	787						
QY	923	GGAGTCTCTTGGTTTTATTTCACTACGCGGAGACGAGGTTGTAAGTGATAGAGAA	982						
DB	788	TCAACCAATTTTAGTCTTTCTCTCGAGAAATAAGACTCATCAAAATTTTGGATTTCTTTTGGATTA	847						

; Patent No. 6696292
 ; GENERAL INFORMATION:
 ; APPLICANT: Stephen M. Allen
 ; APPLICANT: Saverio C. Falco
 ; APPLICANT: Catherine J. Thorpe
 ; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
 ; FILE REFERENCE: BB-1167
 ; CURRENT APPLICATION NUMBER: US/09/720,317A
 ; CURRENT FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: 60/092,833
 ; PRIOR FILING DATE: 14-07-1998
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 1
 ; LENGTH: 2279
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-09-720-317A-1

Query Match 26.8%; Score 655.4; DB 4; Length 2279;
 Best local Similarity 60.4%; Pred. No. 1.6e-171;
 Matches 1120; Conservative 0; Mismatches 726; Indels 9; Gaps 2;

Qy	121	AGTGCACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAGTCTCTAAAGTACTCTTT	180
Db	188	AGTGCACAAGGTGGCGCGCCACCGCGCGAGCACCGCGAGCAAGATGAAGGTGAGGGT	247
Qy	181	GAAGGAGACTTTCTTCCCTGATGACCTTTTGAGGCACTTCAAGAAACAAGCCAGCTTCCAA	240
Db	248	GAAGGAGACCTTTCTTCCCGACGACCGTTCCGCGGCTTCAAGGGGACGCGCGGGAC	307
Qy	241	GAAGTTCACTGGGCTTCAGTCTTCTTCCCATTTTCGAATGGGCTCCCAATAACAC	300
Db	308	GCAGTGGCTCATGCGGTCAGGTACTCTTCCCATCTTCCGACTCGGCTGCGGAGCTACTC	367
Qy	301	CTTTCAGTTCTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCC	360
Db	368	CTTGTGCTCTTCAAGTCCGACCTCTGCGGGGCTCACCATTGCGAGCCTCGCCCATTC	427
Qy	361	TCAGGGCATCAGTTATGCCAAGCTCGCAACCTCCCTCCCAATTTCTGGACTATATTCGAG	420
Db	428	TCAGGGCATTAGCTACGGAGCTGGCAAGCTTGCCTCCCATTAATCGGGCTGTATTCGAG	487
Qy	421	CTTTATACCAACATGATGATATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGGAC	480
Db	488	CTTGTGCGCCGATGTTGATCGCGGTGCTGGGAGCTCCCGTGACCTGGCGGTGGGCC	547
Qy	481	TGTGGCGGTTGGATCGCTTCTGATGGGTCGATGTTGAGTAATGCCGTTGATCCCATGA	540
Db	548	GGTGTGATCTGTCGCTGATCATGGGTCATGGGTCATGCTCGGCGAGCGCTGAGCCCACTGC	607
Qy	541	AGACCCAAAGCTTACCTCCACCTGGCTTTTACAGCTTACATTAATTTGCTGGTGTTCAT	600
Db	608	GGAGCCGACGCTGTTCTGAGCTGGCTTTCACCTCCACCTGTTCCGCGGGCTGGTGCA	667
Qy	601	GGTGGCTTGGGCTGTTGATGGGTTGATGGGTTGATGGGTTGATGGGTTGATGGGTTGAT	660
Db	668	GGCTCCCTGGGATCTCAGGCTCGGCTTCTGATCGGCTTCTGATCGGCTTCTGATCGGCT	727
Qy	661	AATAGGTTTCATGGGAGGACGCGGTTGCTGCTGAGGTTGCTGAGGTTGCTGAGGTTGCT	720
Db	728	GGTGGGTTTCATGGCGCGCGCCGATCATCGTGGGCTGCGACACTCAAGGGGCTGCT	787
Qy	721	TGGCTTGGAGCATTTACCATCGAGCTGATATCATATCATGATGATGGCTCTGTTTTCAC	780
Db	788	GGGATCGTCCATTTACACACCGAGATGGGATCGTCCGAGTATGAGGCTTCGTTCTCCA	847
Qy	781	CCAAACTCATGAGTGGAGGTGGAAAGTGTGTTAGGATGTGTTTCAATTTTCTTCTCT	840
Db	848	CCACACGAGGAGTGTGCTGGGAGAGCATCTCATGGGCTGCTGCTTCTCTCTTCTCT	907
Qy	841	CTTAGCACAGATCTTACGAAAAAACAACCCAGGTTTTTTTGGTGTGACCAATGGC	900

Db	908	GCTGTCGGCGAGGCATGTGAGCATCAGATGGCCAAAGCTTTTCTGGGTTTCGGCGTGGC	967
Qy	901	GCCATTGACGTCGCTTATATTTGGAAAGTCTCTTTGGTTTATTTTCACTCACGCCGAGAAGCA	960
Db	968	GCCCTGGCATCGTACCACCTCTCGACGCTGCTTTTCTCTTCAAGCTCAGAACCA	1027
Qy	961	CGGTGTTGAAGTGTAGGAGAACTGAAGAGGGTTTGAATCCACCATCACTCACAATCT	1020
Db	1028	TGGCATCAGCATCATTTGGGAGCTCAAGTGGCGCTGAATCGCCCTCTGCTGGGACAGCT	1087
Qy	1021	GATATTGTTGTCGCTTACATGACTACAGCTGTCAAAACTGGCATTTGCTGTTGGCATCAT	1080
Db	1088	CTGTTTGACACGCGCTATTTAGGCTCACCATGAAGCTGGGCTGTGACCGGAATCAT	1147
Qy	1081	ATCACTTGGGAAAGGAATAGCAGTAGGAAGAAGCTTTTGCATATGTATAAAATTAACAATAT	1140
Db	1148	CTCACTGACGGAAGGAATAGCGTGTGTAGAACAATTTGCTCTCACTCAAGGACTACCAT	1207
Qy	1141	TGATGGAAACAAAGATGATAGCTATTTGGGACCATGAAGCTAGTTGGTTCTTTTCACTC	1200
Db	1208	AGATGGAAACAAAGGATGATGGCCATAGGGTTGATGAATGTTGTTGGGTCCTGCACATC	1267
Qy	1201	TTGCTTACCTCACAAACAGGACCAATTTTCGGGTTTCGGCTGTGAACCTATAACGCTGATGCAA	1260
Db	1268	ATGCTAGTAACAAACAGGTGCTTCTCCGCTCTGCTGTAAACACACACCGCGCTGCAA	1327
Qy	1261	GACGAGCTTCCAAACATTAATGTCACCTTGCAGTAATGTTGACATTTGTTATTCCTGAC	1320
Db	1328	GACTGCCATGTCCAAACGTCATCATGCGCTGACTGTGATGGTCACGCTGCTTCTCTCAT	1387
Qy	1321	ACCTTGTTCATTAACCTCCCTGGTGTGCTATCAGCTATTTATCGTATCTCGAATGCT	1380
Db	1388	GCCACTGTTGCTGTACACACCCAAACGTTGTCTCGAGCGATCATCATCGCCGCTGAT	1447
Qy	1381	TGCACTCATAGATTAATGAAGCAGCCATCCATCTATTTAAAGTTGACAAATTTGACATTTGT	1440
Db	1448	CGGCTGTGATGATTTCCCGCGGTGTACACATCTGGAAGATGGAAGATGATGATTTCT	1507
Qy	1441	GCTGTGATGAGTGATATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCTAT	1500
Db	1508	GGTGTGGTTTGGCGCTTTGCGCGCTCATCTTTCATCTCAGTCCAAAGAGGCTTGGAT	1567
Qy	1501	AGCTATTGTAATCTGTACTTCTGGGTACTCTTATTTATTTGAAGGCAAGGACATTCGT	1560
Db	1568	AGCGTGTGATATCTATTTAGGCTGTGATGACAGTCAACAGGCGGAGATGATGCT	1627
Qy	1561	TTTGGGCAACATTCCTGTATATACCGAAATGTTGAGCACTATCAAAATGCAAA	1620
Db	1628	TCAGGGAACATCAGGGGACTGATTTTACAGAGACCTGCATCATCAAGAGGCGCCA	1687
Qy	1621	ACATGTTCTGGAAATCTAATTTCTAGAGATTTGATGCACCAATTTTACTTTGCCAATGCCAG	1680
Db	1688	AAGATTTCTGGGTTCTTGTATCTTGGCTTGAAGCACCAGTAACTTGCSCCACTCCAA	1747
Qy	1681	CTATTGAAGAAAGGATCAAGGTGGATTTGATGAAGAAAGAAAGAAATTAAGCTAC	1740
Db	1748	CTACCTGAATGAAGGATTAAGAGTGGAT-----AGAGGAAGAAATCTTTTGAACAGGA	1801
Qy	1741	AGGGGAGACTAGTTTCAGTATGTTAATTTGATATGAGTGTGTTGGAAACATTTGATAC	1800
Db	1802	TAAACATCTGAACCTCCATTTCAATCTTGGATCTGTGAGCTGTTCTTGCATTTGACAC	1861
Qy	1801	AAGTGAATAAGTATCTCTTGAAGAGGTGAAGAGATTTACAGAGAGAGAGAGCTACAGCT	1860
Db	1862	AAGTGCATAGCGTTCCTCATTTGACATAAAGAAATCAATAGAGAAAGCTGCTCTGGAGCT	1921
Qy	1861	TGTTTGGTCAATCTGTGAAGTGAAGTGAAGAACTGAAACAAATCGA---AGTTCCA	1917
Db	1922	TGTGCTGTCAATCCAACTGGAGAGTCAATGGAGAAATAACAACGTCGCAACCGAGGCTGA	1981
Qy	1918	AATCATTTAGGGAAGAAATGGATCTATCTGATGTTGAAGGCGCTTGGAGCA	1972
Db	1982	AACTATTTTAGGCCAGATTTGCTTGTATCTGACCACTGGCGAAGCAATGCTTTCA	2036

RESULT 5

US-09-720-317A-19
; Sequence 19, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 2311
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-720-317A-19

Query Match 26.1%; Score 640.4; DB 4; Length 2311;
Best Local Similarity 59.4%; Pred. No. 2.3e-167;
Matches 1104; Conservative 0; Mismatches 751; Indels 3; Gaps 1;

RESULT 5																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									</
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Db 1927 CTCATTCTGTCGAATCCTGGTTCGTCGTATAGAAAACTGCAAGCGTCGAAGCTCACC 1986
Qy 1919 AATCAATTAGGAAGAAATGATCTATCTGACTTTGAAAGAGCGCTGGAGCATGCA 1976
Db 1987 GAGCATTGGAAGCAGCAATATATCTCGCGTCTCTGACGCTGTGCGATTCTGTA 2044

RESULT 6
US-09-720-317A-7
; Sequence 7, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Catherine J. Thorpe
; APPLICANT: Saverio C. Falco
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Helianthus tuberosus
US-09-720-317A-7

Query Match 13.2%; Score 324; DB 4; Length 780;
Best Local Similarity 70.6%; Pred. No. 1.2e-79;
Matches 432; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

Qy 1400 GCAGCCATCCATCTATTAAAGTTGACAAATTTGACTTTGTGTGTCGATGAGTGATAC 1459
Db 17 GCSCGATTACCTCTCGACACATAGACAAATTCGACTTTGTGTGTCGATGAGTGATAC 76
Qy 1460 ATTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCTAGTATGTAATCTGTA 1519
Db 77 TTTGGTGTCTTTGGAGTGTGAAATTTGGATTAGTATCGCGTCCGATTTGCTGTTG 136
Qy 1520 CTTGGGTACTCTATTATTGCAAGGCCAAGGACATTCGTTTGGCACAATTCCAAAT 1579
Db 137 CTTAGGGTACTCTCTATTGTTCTGAGGCCAAGAACATCGAGCTAGTCTCATACCCGAT 196
Qy 1580 TCTGTGATATACCGAAATTTGAGCACTATCAAAATGCAAAACATGTTCTCGAATGCTA 1639
Db 197 TCCACTATTATAGAAGTATGATCAATACCAAAATCGGAAGCGTTCCAGGAATCTTG 256
Qy 1640 ATTCTAGAGATTGATGCCAAATTTACTTTGCCAATGCCAGCTATTTAAGAGAAAGATC 1699
Db 257 ATACTTCAAATCGAAGCACCTATTACTTTGCTAACTCTAGCTACTTTGAGGGAAGGATT 316
Qy 1700 ACAAGGTGGAATGATGAAGAGAGAAATTAAGCTACAGGGAGACTAGTTTGCAG 1759
Db 317 GTGAGATGGGTGATGAAGAGAGAGATAGGTGAAGTCTTTAAGAGAGATGACTTGCAA 376
Qy 1760 TATGTTAATATTGATATGAGTGTGTTGGAACATTTGATACAAATGGAATGAATGCTT 1819
Db 377 TATGTCAATCTTGCAATTGAGTGTGTTGGAATATTGATACAAATGGAATGAATGCTT 436
Qy 1820 GAGAGGTGAAGAAGATTACAGAGAGAGAGATCAGCTTGTGTTGGTCAATCTCTGTA 1879
Db 437 GGAGAAGTTAAAGAGTTATGGAAGAGAGAGGCTAAAGTTGGTTTATAGCGAATCCGGGC 496
Qy 1880 AGTGAAGTGTGAAGAATCGAATCGAAGTTCCAAATCATTTTAGGGAAGAATGG 1939
Db 497 GGAGAGTAAAGAAGATGAACAAAGCGAAGTTGATAGAGTGATCGGGGGAAGATGG 556
Qy 1940 ATCTATCTGACTGTGTAAGAGCGCTTTGGAGCATGCAACTTTCAATCTACCTGCAAGCAA 1999
Db 557 ATATACTACAGTGGGAGAGCGGTTGGAGCGTGCAACTTTATGCTTACTACTACAG 616

Qy 2000 ACGAACCCAAAG 2011
Db 617 AACGCCCAAAAG 628

RESULT 7
US-09-720-317A-5
; Sequence 5, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-5

Query Match 8.8%; Score 216.2; DB 4; Length 1240;
Best Local Similarity 53.4%; Pred. No. 1.4e-49;
Matches 477; Conservative 0; Mismatches 413; Indels 3; Gaps 1;

Qy 1085 CTTGCGAAGGAATAGCAGTAGGAAGAAGCTTTGGCAATGTATAAAAATTAACAATTTGAT 1144
Db 9 CTCACGGAAGCTATCGCGTTGGCCGATCTTTCCCTCCGTAAGAGGGTACAGACTCGAC 68
Qy 1145 GGCACAAAGAGATGATAGCTATTGGGACCATGAAAGTAGTTGGTTCTTTTCACTCTTGC 1204
Db 69 GGCACAAAGAGATGCTGGCCATGCGGTCTTCCAAGCTTGTCTGTCTCTCTCTCGTGC 128
Qy 1205 TACCTCACAACAGAGACATTTTTCGCTTGGCTGTGAACCTATATAACGCTGATCAAGACA 1264
Db 129 TATGTGGCAACAGGTTCTGTTCTCCGAAACGAGTGAACCTTCAGCGGGGGCCAGGTGC 188
Qy 1265 GCAGCTTCCAAACATTAATATGTCATTTGCAGTAATTTGACATTTGATTTTCCTGACACC 1324
Db 189 ACCGTTTCAAACATCTGTCATGTCATCACCGTGTTCCTCACCTGGAGCTGTTTCATGAAG 248
Qy 1325 TTGTTCCATTAACCTCCCTGGTGGTCTATCAGCTATTATCGTATCTGCAATGCTTGA 1384
Db 249 CTCCTCTACTACACGCCCATGCGGTGCTCGCCTCCATCATCTCTCGGCTCTTCGGGA 308
Qy 1385 CTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTTGGTG 1444
Db 309 CTGATCGACATCAAGAGAGCGCTGCAGCATATGGAAGATCGACAAGATGATTTCTCACC 368
Qy 1445 TGCAAGTGCATACATTCGCGTGGTCTTTGGCAGTGTGAAATTCGCTTAGTATAGCT 1504
Db 369 TGCCTCGTGCCTTTGTTGGCGTCTGTTGGTGGTGGAGATTTGGCTTGCAGTTGCA 428
Qy 1505 ATTGTAATATCTACTCTCGGTACTTCTATTATTGCAAGGCCAAGGACATTCGTTTTCG 1564
Db 429 CTTGGCATTTCTTCCGAAAGATCATCATACAGTCGCTTCGGCTTCAGGTGGAGATCTT 488
Qy 1565 GGCACATTTCCAAATTTCTGTGATATACCGAAATTTGAGCACTATCAAAATGCAAAACAT 1624
Db 489 GGCAGCTACAGGACAGATATCTTTCGAGCGTCAGGAGTACCTCTGAGCTGCCTA 548
Qy 1625 GTTCTCGAATGCTAATTTCTAGAGATGATGACCAATTT---TACTTTGCCAATGACG 1681
Db 549 ACTCCGACTGACTGCTATAGCGTCGACACATCTCTCTCTCTCTCATCAACGCACT 608
Qy 1682 TATTTAAGAGAAAGGATCAAGGTTGATGAGAGAGAAAGAAATTAAGCTACA 1741

Db 609 TCCGTCACAAAGAAAGGATCACAGAGTGGGTTTGGGAAGGAGTGGAGACCTCAAAATGGAAAA 668
Qy 1742 GGGGAGACTAGTTTGGCAGTATGTTATAATTGATATGAGTCTGTTGGAAACATTTGATACA 1801
Db 669 GCAGGGAGAGGATACAGCAGTTGCTTGTATGTCAAGTGTGTTAAACATCGACACT 728
Qy 1802 AGTGGATAAGTAGTCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGACTACAGCTT 1861
Db 729 TCAGGACTCACTGCACTGGGAAGAAATACACAAGGAGTTGGTGCTCTTGGCTTACAGATG 788
Qy 1862 GTTTTGGTCAATCTGTAAGTGAAGTGAAGAAGAACTGAACAAATCGAAGTTCCAAAAT 1921
Db 789 GCTATAGCCAGTCCGGGATGGAAGGCAGTTTCAAGAAGATGAAGTGTCAAGAGTGGTGGAC 848
Qy 1922 CATTTAGGGAAGAAATGGAATCTATCTGACTGTTTGAAGAGGCCCTGTGAGCATG 1974
Db 849 AGGTAGGACAGGACTGGATCTTCATGACAGTAGGTGAAGCGGTGGAGGCCTG 901

RESULT 8
US-09-720-317A-21
; Sequence 21, Application US/09720317A
; Patent No. 6698292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720, 317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-720-317A-21

Query Match 8.4%; Score 205.6; DB 4; Length 2022;
Best Local Similarity 47.0%; Pred. No. 1.7e-46;
Matches 791; Conservative 0; Mismatches 869; Indels 24; Gaps 4;
Qy 306 AGTCTTTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTGGCCATTCCTCAGS 365
Db 103 AGGACTTCAGGCGGACCTCGCCCGGCGATCACTGTGCGCGTCATGCTTGCCTCAGG 162
Qy 366 GCATCAGTTATGCAAGCTCGCCAACTCCCTCCAATTTCTTGACATATATTCGAGCTTTA 425
Db 163 CAATGTCATATGCAAGCTGGCTGGGCTTCACCCAAATTTATGGCTCTACACAGGCTTG 222
Qy 426 TACCACATTTGATATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAGCTGTGG 485
Db 223 TCCCACTATTGTCTACGCGGAATTTTGGGTCTCTCAGCAAAATAGCAGTAGGTCCAGTGG 282
Qy 486 CGGTTCGATCGCTTCGTAGGGTTTCGATGTTGAGTAATCGGTTGATCCCAATGAGACC 545
Db 283 CACTGTCTCTGTGATGTCOAATGTTCTTGGGGGTATAGTTAATTCATCTAGTG--- 339
Qy 546 CAAAGCTTTACCTCCACCTGGCTTTTCACAGCTACATTAATTTGCTGTGTTTTTCAGGCTG 605
Db 340 ---AGCTGATACGGAATAGCCATATATTTGSCATTCATGTTGGAATACTGGAATGCT 396
Qy 606 CCTTGGGTCTGTTAGTGGGTGATCGTGGATTTTCTGTCATATGCAATGCAACCATATAG 665
Db 397 TGATGGCAATTCGTAAGACTTTGGCTGCTTATTCGTTTCAATAGCCATTCGTGAATATCTG 456
Qy 666 GGTTCATGGGAGGAGCAGCAGCTGTGTCGAGCAACTAAATCGATTCGATTCGTGGCC 725
Db 457 GATTCTACAGCTTCGGGCCATCGTAATTTGGTTTGTGCCAAATCAAGTATTTCTTGGG--- 514

Qy 726 TTGAGCATTTCCACCCATGAGCTGATATCATATCAGTGTGCGCTCTGTTTTCCACCAA 785
Db 515 -TTACAGTGTACAGAAGTAGCAAAATATATACCCTTATTGAGAGTATAAATTGCTGGAA 573
Qy 786 CTCATGAGTGGAGGTGGGAAAGTGTGTGTAGGATGTCTTCATTTTCTCTCTCTTA 845
Db 574 TAGATCAGTTCTCTCGGCTCCATTTGTAAATGGGATCAGCGTTTCTTGTATTCTTCTAA 633
Qy 846 GCACAGATACCTTCAGCAAAAACGACCAAGGTTTTTTTTGGGTGTGAGCAATGGCGCCAT 905
Db 634 TAATGAAAAGCTAGGGAACAAATAAATAATACGTTTCTCTGAGAGTCTCTGTTCCAC 693
Qy 906 TGACGTCCGTATATTTGGGAAGTCTCTTGGTTTATTTTCACTCACGCGGAGAGACGCTG 965
Db 694 TAACAGCTGTCTTCTTGGAACTTGTGTGAAATTTTCCGTCC-----AAGTCCCA 747
Qy 966 TTGAAGTGTAGAGAGAACTGAAAGAGGTTTGAATCCACCATCACTCAAAATCTGGTAT 1025
Db 748 TATCAGTGTGTAGTGAATACCCCAAGGCTTTCCCAAGTTTCTCCATTCCTCGAGGATTTG 807
Qy 1026 TTGTGTGCGCTTACATGACTACAGCTGTCAAACTGGCAATTCGTTGGCATCATATCAC 1085
Db 808 AACATCTGATGCCCTAATGCCAACTGCAATACTTATCACTGCTGTTG-----CTA 858
Qy 1086 TTGCGGAAGGAATAGCAGTAGGAAGGCTTTTGCATATGTATAAAATTAACAATATTGATG 1145
Db 859 TTTTGGAGTCTCTTGGGATTGCTAAAGCGTTAGCTGCGAAGAATGTTTATGATGGACT 918
Qy 1146 GCAACAAAGATGATAGCTATTGGGACCATGAAGTAGTGTGGTTCTTTCACCTCTTGCT 1205
Db 919 CAAACAAAGAGTATTATTGGCCTTGGCTTATCAAAATATATGCGGTTTCATTTCTCTGCA 978
Qy 1206 ACCTCACAAACAGGACCATTTTCGCTTGGCTGTGAACCTATAACGCTGATGCAAGACAG 1265
Db 979 ATCTGCTACAGGCTCTTTTCTAGTCTGCTGTGAATCATGAAAGCGGGGCAAGACTG 1038
Qy 1266 CAGCTTCCAACATTAATGTCACTTGGCAGTAATGTGACATTTGTTATTCCTGACACCT 1325
Db 1039 GATTATCAGGAATATAATGGGCATAAATAATTTGCAGTGTCTCTTGTATTATGACACCAT 1098
Qy 1326 TGTTCCATTACACTCCCTGTTGGTGTCTATCAGCTATTATTCGTATCTGCAATGCTTGGAC 1385
Db 1099 TATTTACTGATATACCTCAGTGTGCAATTTGGCTGSCCAATTTGTGATTTCTGCTCAGCTGGCC 1158
Qy 1386 TCATAGATTATCAAGCAGCCATCCATCTAATTTAAAGTTTGACAAATTTGACTTTTGGTGT 1445
Db 1159 TGGTAGATTGAAAGAGGCCATCTTCTGTGGGTTATTTGATGAAGAGGATTTCTTCTGT 1218
Qy 1446 GCATGAGTGCAATCAATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTTAGTCATAGCTA 1505
Db 1219 GGGCGATGACATTTACTACAACCTTAACTTTTGGCAATTTGAGATTGCTGTCTTGTGGGG 1278
Qy 1506 TTGTAATATCTGACTCTCGGTACTTCTATTATTATTCAGAGGCCAAGGACATTTGCTTTTGG 1565
Db 1279 TCGGTTTTTCGTGTCATTTGTGATCCCATGAATTCGAAATCCGCATATAGCTGTTTTGG 1338
Qy 1566 GCAACATTCCAAATTTCTGTGATATACCGAAATGTTTGAGCACTATCAAAATGCAAAACATG 1625
Db 1339 GCGGTTTGGCTGGCACCACTGTGTACAGGAATACATTTGCAGTACCTCTGAGGCTTATACAT 1398
Qy 1626 TTCTTGGAAATGCTAATCTAGAGATTGATGCAACCAATTTACTTTGCCAATGCCAGCTATT 1685
Db 1399 ACAACGGGATTTGTTGTGCTGTGATGACCACTCTACTTTGCTAACAATAGTTTACA 1458
Qy 1686 TAAAGAAAGGATCACAGGTGGATTGATGAAGAGAAAGAAATTAAGACTACAGGGG 1745
Db 1459 TAAAGGACAGGTTGCTGAGTATGAGCTCAAACTCCCAATTCAAACCGTGGACCTGATG 1518
Qy 1746 AGACTAGTTTTCAGTATGTTATAATTTGATATCAGTGTCTTGGAAACATTTGATACAAGTG 1805
Db 1519 TTGGAAAGGTTGACTTTTGTGATCTCTCGAGATGTCCTGCTGTTACATACATCGACTCGAGCG 1578
Qy 1806 GAATAAGTATGCTTGAAGAGGTGAAGAAATTTACAGAGAGAAAGAGAGCTACAGCTTGT 1865

Db 1579 CTGTTCAAGCTCTCAAGGACCTGACCAAGACCTGACCAAGATACAAAGACGCGGACATCCAGATTGCTA 1638
Qy 1866 TGGTCAATCTCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCAAAATCATT 1925
Db 1639 TAGCGAATCTTAACCGGAGGTGACCTATTGCTGTCAAGAGCGGCATCATCGACATGA 1698
Qy 1926 TAGGGAAGAATGATCTATCTGACTGTGTAAGAGCGCGTGGAGCATGCAACTTCAATC 1985
Db 1699 TTGGCGCAGGCTGCTGTTTCGTCGAGTGCACGACGCGGTGCAAGTATGCTCCAGCATG 1758
Qy 1986 TACG 1989
Db 1759 TCGG 1762

RESULT 9
US-09-720-317A-13
; Sequence 13, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-720-317A-13

Query Match 4.8%; Score 118; DB 4; Length 493;
Best Local Similarity 55.0%; Pred. No. 1.7e-22;
Matches 252; Conservative 0; Mismatches 205; Indels 1; Gaps 1;
Qy 816 TAGGATGTCTTCATTTTCTCTCTCTAGCACAGATCTTCAGCAAAAAGACCAA 875
Db 36 TCGGATGCTCTTCTCTCATATTCATCTCCACACAGGTTTCATCGGGAGAGGTACAA 95
Qy 876 GGTGTTTTGGGTGTGAGCAATGCGGCATTCAGTCCGTTATATTTGGGAAGTCTCTGG 935
Db 96 AGCTGTTCTGGGTGTGAGCGATCTCGGCTCTGCTGCTGCTCTCTCTCTCTCTCTCT 155
Qy 936 TTTATTTCTACGCGCGAGAAGCAGCGGTGTTGAAGTGTAGGAGAACTGAAGAAGGGTT 995
Db 156 TCTACGGGCAAGGGGTGACAGGCGCGGTCAAGATCATCCAGAGGTGCAACGCGGCG 215
Qy 996 TGAATCCACCATCTACT-CACAAATCTGTTATTTGTGTGCGCTTACATGACTACAGTGT 1054
Db 216 TAAACCCCAAGTCTCGTGGAGAGATACACCTCAACGCGGCGGCACACACAGGAGT 275
Qy 1055 AAAATCGGCATGTCGTTGGCATCATATCTCTGCGGAAGAAATAGCAGTAGGAAGAAGC 1114
Db 276 CAAGATCGCGTCTATCTGCGCATCATCGCCCTCAGGAAAGCTATCGCCGTGGCCGATCT 335
Qy 1115 TTTGCAATGTATAAAATTAACAATTTATGATGGCAACAAAGAGATGATGATTTGGGACC 1174
Db 336 TTGCGCTCCGTAAGAGGGTACAGACTCGACGGCAACAAAGAGATGCTGGCCATGGGTTT 395
Qy 1175 ATGAACTAGTGTGTTCTTCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1234
Db 396 TCCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
Qy 1235 GCTGTGAACATATAACGCTGGATGCAACACAGCAGCTTC 1272
Db 456 GCAGTGAACCTTACGCGCGGGGCGGCGGCTGACCGGTTTC 493

RESULT 10
US-09-720-317A-9
; Sequence 9, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR FILING DATE: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (398)
; OTHER INFORMATION: n = A, C, G or T
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (483)
; OTHER INFORMATION: n = A, C, G or T
US-09-720-317A-9

Query Match 4.8%; Score 117.4; DB 4; Length 484;
Best Local Similarity 55.7%; Pred. No. 2.5e-22;
Matches 265; Conservative 0; Mismatches 207; Indels 4; Gaps 2;
Qy 717 TTCTTGGCCTTGAGCATTTTCAACCCATGGAGCTGATATCATATCAGTATGCGCTCTGTTT 776
Db 2 TTACAGGCATCAAGAGCTTTTACAAAGAAACCGATATAATTTCCGTGATGAGCTGAAGTC 61
Qy 777 TCACCCAAACTCATGA---GTGGAGGTGGAAAGTCTGTGTTAGGATGTCTTCATTT 833
Db 62 CAACAGAGGGGNTAACAGGTGGAATTTGGCAAACTATTGTGATGGCATAACTTTCTTTG 121
Qy 834 TCTTCTCTTACACAGATATTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTGTCAG 893
Db 122 CATTCCTTCTGCTTGCCAAAGTACATTGGAAAGAAAGAAATAGGAAGTTCTTCTGGGTG 181
Qy 894 CAATGGCGCATTCAGCTCGTTATATTGGGAAGTCTTGTGTTTATTTTCACTCAGCGCG 953
Db 182 CTATTGCTCTATAACTTTCAGTTATTTTGGCAACCCCTTTTGTGTTTCATTACTCGTGTG 241
Qy 954 AGAAGCACGCTGTGAAGTGTAGGAGAACTGAAGAAGGTTTTGAATCCACCATCACTCA 1013
Db 242 ACAGCAAGGTGTTTCAAGTTGTTAACCATCAAAGGGGCAATAACCATCATCATAGTCC 301
Qy 1014 CAAATCTGTTATTGTGTGCGCTTACATGACTACAGCTGTCAAAGCTGGCATTTGCTGTG 1073
Db 302 ACAAAATTTATTCTACTGCTGCTTCAAGTTCGAAAGGTTTCAAGATCGGTGTCATTTCCG 361
Qy 1074 GATCATATCATCTGCGGAAGGAATAGCAGTAGGAAGAGAGCTTTTGCATGTATATAAAT 1133
Db 362 CCATGATCGGTTTAAACGGAAGCTGTGGCAATTTGGGANGAGCTTTTGTGCTCTGAAGACT 421
Qy 1134 ACAATATTGTCGCAACAGAGATGATAGCTATTTCGGACCATGAACGAGTAGTTGGT 1189
Db 422 ATCAATTAGATTGGAACAAGGAGATG-GAACACTTGGAACTATGAACATACAAAGGT 476

RESULT 11
US-09-785-381-4
; Sequence 4, Application US/09785381
; Patent No. 6602992

; Sequence 5627, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5627
 ; LENGTH: 4927
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-5627

Query Match 3.4%; Score 82.6; DB 4; Length 4927;
 Best Local Similarity 47.0%; Pred. No. 4.3e-12;
 Matches 290; Conservative 0; Mismatches 324; Indels 3; Gaps 1;

QY	1087	TGCGGAAGCAATAGCAGTAGGGAAGAGCTTTTGCATATGTAATAAATTTACAATATTGATGG	1146
DB	1307	TGCTATTGAGTGTCTGATAGGAAAGATATATGCCCAAGATGATGATACCAATCGATGG	1366
QY	1147	CAACAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTCTTCACTCTTTGCTA	1206
DB	1367	GAACCAAGCAATCATTTGCGCTTTGGGATCAGCAACATCTCTCAGGATTTCTCTCTGTTT	1426
QY	1207	CCTCAACAAGACCAATTTTTCGGTTCGGCTGTGAATCTATTAACGCTGGATGCAAGACAGC	1266
DB	1427	TGTGGCCACCATCTGCTCTTTTCCCGCACCGCCGCTCAGGAGAGCACTGGAGGAAAGACACA	1486
QY	1267	AGCTTCCAAACATTAATAATGTCACTTGCAGTAATGTGATGATGATGATGATGATGATGATG	1326
DB	1487	GGTTGCTGGCATCACTCTGCTGCGATTTGATGATGATGATGATGATGATGATGATGATGATG	1546
QY	1327	GTTCCTCAATACCTCCCTGGTGTGCTATCAGCTATTATCGTATATCGTCAATGCTTTGGACT	1386
DB	1547	TCGTGAACCTTTGCAGAAAGTCGGTCTTGGCAGCTGTGTAAATTGCAACCTGAAAGGAT	1606
QY	1387	CATAGATTATGAAGCAGCCAT--CCATCTATTTAAGGTTGACAAATTTGACTTTTGTGCT	1443
DB	1607	GTATTATGACAGTGTGTGACATTCCTCGTCTGGGAGACAGATAAGATGATGCTGTTAT	1666
QY	1444	GTGCATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGC	1503
DB	1667	CTGGGTGTTTACGTGTATAGTGTCCATCATCTCTGGGCTGGATCTCGGTTTACTAGCTGG	1726
QY	1504	TATTGTAATCTGTACTTTCGGGTACTTATTTATTTGCAAGGCCAAGACATTCGTTTTT	1563
DB	1727	CTTATATTGGACCTGTGACTGTGTCTGAGAGTTTCTGTTCTCTTCTGGAATGGCCT	1786
QY	1564	GGGCAACATTTCCAAATTTCTGTGATATACCGAATTTTGGACACTTCAAAATGCAAAACA	1623
DB	1787	TGGAAGCATCCCTAGCACAGATATCTACAAAAGTACCAGAAATTTACAAAACATTTGAAGA	1846
QY	1624	TGTTCTCGGAATGCTAAATTTCTAGAGATGATGCACCAATTTTACTTTTGCATATGCCAGCTA	1683
DB	1847	ACCTCAAGAGTGAAGATTTCTTAGATTTTCCAGTCTTATTTTCTATGTCATGTCGATGG	1906
QY	1684	TTTAAGAGAAAGATCA	1700
DB	1907	TTTTAAAAATGTATCA	1923

; GENERAL INFORMATION:
 ; APPLICANT: DALLOS, Peter
 ; APPLICANT: ZHENG, Jing
 ; APPLICANT: MADISON, Laird
 ; TITLE OF INVENTION: A MAMMALIAN PRESTIN
 ; FILE REFERENCE: 0290-37U1
 ; CURRENT APPLICATION NUMBER: US/09/785,381
 ; CURRENT FILING DATE: 2001-02-16
 ; PRIOR APPLICATION NUMBER: US 60/183,461
 ; PRIOR FILING DATE: 2000-02-18
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 2441
 ; TYPE: DNA
 ; ORGANISM: Mus sp.
 ; US-09-785-381-4

Query Match 3.7%; Score 90.4; DB 4; Length 2441;
 Best Local Similarity 46.9%; Pred. No. 2e-14;
 Matches 317; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

QY	1061	GGCAATGTCTGGCATCATATCACCTTGGGAAGGAATAGCAGTAGGGAAGCTTTGCCA	1120
DB	1233	GCATTTGCCATCGCCATCGTTGGATTTTCAGTGACGATCTCATCGGCCAAACCTTTGCCA	1292
QY	1121	ATGTATAAAAAATTTACAATATTGATGGCAACAAGAGATGATAGCTATTGGGACCATGAAC	1180
DB	1293	AATAAGCATGGCTACCAGGTTGATGGCAATCAGGAGCTCATTTGCCCTTTGGGATATGCAAC	1352
QY	1181	GTAGTTGGTCTTTACCTCTTGTCTACCTACCAACAGGACCAATTTTCGGTTCGGCTGTG	1240
DB	1353	TCCATTTGGATCTCTCTTCCAAACCTTTCTCGATTTCTGCTCTGCTCTCGAAGCCCTTGT	1412
QY	1241	AACTATAACGCTGGATGCAAGACAGCAGCTTCCAAACATTAATAATGTCACTTGCAGTAATG	1300
DB	1413	CAGGAAGAACTGGAGGAAACAACAGCTTGAGGTTGTTGGCCCTCGTTGAATGATCTG	1472
QY	1301	TTGACATTTGTAATTCCTGACACCCCTTTGTTCCATTACACTCCCTCGGTGGTGTATCAGCT	1360
DB	1473	TTGGTCATATTAGCCACCGGATTCCTCTTTGAGTCGTTAGCCAGGCTGTCTCTTCGCCC	1532
QY	1361	ATTATCGTATCGCAATGCTTGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAG	1420
DB	1533	ATTGTGATCGTCAACCTGAAGGAATGTTTCATGCAAGTCTCAGACCTCGCCTTTTGTGG	1592
QY	1421	GTGGA---CAAAATTTGCTGTTGGTGTGATGATGATGATGATGATGATGATGATGATGATG	1477
DB	1593	AGAACAGCAAAATAGAGCTGACCATCTGGCTGACCACTTTGTGTCTCTCCTGTTCCTC	1652
QY	1478	AGTGTGAAATTTGGCTTAGTCATAGCTATTGTAATATCTGTAATTTTCGGGTACTTCTATT	1537
DB	1653	GGCTTGACCTAGGACTGATTACCGCGGTGATCATTTGCTCTGCTCACAAGTATTATAGA	1712
QY	1538	ATTGCAAGGCCAAGGACATTCGTTTGGCAACATTTCCAAATCTGTGATATATACGGAAT	1597
DB	1713	ACACAGAGTCCAAAGCTACAAAGTCTCTGGGGAGCTCCCTGACAGCGATGTGTACATTGAC	1772
QY	1598	GTGAGCACTATCAAAATGCCAAACATGTTCTCTGGAATGCTAAATTTCTAGAGATTGATGCA	1657
DB	1773	ATAGATGCAATGAGAGAGTGAAGAAATTTCTTGGAAATTAATAATTTCCAAATTAATGCC	1832
QY	1658	CCAATTTACTTTGCCAATGCCAGCTATTATTAAGAGAAAGGATACCAAGGTGGATGATGAA	1717
DB	1833	CCAATTTACTATGCAATAGGACTTGTATAGACGCGCTTTTAAAAAAGAAAGACTGGAGTA	1892
QY	1718	GAGAGAGAAAGATTA	1733
DB	1893	AACCCAGCACTCATTA	1908


```
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300.6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMMU
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703)836-9300
/ TELEFAX: (703)683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7218 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ CLONE: pTZgpt-Fls
/ US-08-232-463-14

Query Match      3.1%; Score 76.4; DB 1; Length 7218;
Best Local Similarity 6.3%; Pred. No. 2.8e-10;
Matches 26; Conservative 235; Mismatches 151; Indels 0; Gaps 0;

Qy 1672 CAATGCCAGCTATTTAAGAGAAAGGATCAACAAGGTGGATTGATGAAGAAGAAAGAAT 1731
Db ||||| || ||||| ||||| ||||| : : : : : : : : : : : :
Qy 1471 CTATGCAAGTAGTTAAAGAGATAGAGAAATTTGGTACRRRRRRRRRRRRRRRRRR 1412
Db

Qy 1732 TAAAGCTACAGGGGAGACTAGTTTGCAGTATGTATAATTGATGATGCTGTTGAAA 1791
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1411 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1352
Db

Qy 1792 CATTGATACAAAGTCGAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGA 1851
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1351 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1292
Db

Qy 1852 GCTACAGCTTGTGTTGTCATCTGTAAGTGAAGTGAAGAACTGAACAAATCGAA 1911
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1291 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1232
Db

Qy 1912 GTTCCAAAATCATTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGCCGCTTGAGC 1971
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1231 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1172
Db

Qy 1972 ATGCAACTTCAATCTAGTCAAGCAAAACCAACCAAGAAAGATCAACAGAGCGTTG 2031
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1171 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1112
Db

Qy 2032 GAACAATGTGTGACTAGTCAATATGCCAAAGATATTTCTAAATAACTCAAAA 2083
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1111 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1060
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: August 31, 2005, 08:53:07
Job time : 435 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 03:14:21 ; Search time 2765 Seconds
(without alignments)
5795.308 Million cell updates/sec

Title: US-10-762-049-17

Perfect score: 2449

Sequence: 1 gcacgagctagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1521.6	62.1	2008	18	US-10-424-599-131116
2	1018.2	41.6	1977	9	US-09-938-842A-1502
3	1018.2	41.6	1977	11	US-09-938-842A-1502
4	846	34.5	2372	19	US-10-437-963-38184
5	790.2	32.3	1518	18	US-10-424-599-131117
6	677.4	27.7	2727	20	US-10-425-115-50601
7	670.4	27.4	4390	19	US-10-437-963-77012

8	658.6	26.9	2366	18	US-10-424-599-65103	Sequence 65103, A
9	650.6	26.6	2651	20	US-10-425-115-182207	Sequence 182207, A
10	647.2	26.4	2780	19	US-10-437-963-77005	Sequence 77005, A
11	600.6	24.5	2001	19	US-10-437-963-95853	Sequence 95853, A
12	571.4	23.3	2429	19	US-10-437-963-73410	Sequence 73410, A
13	567	23.2	2656	18	US-10-425-114-15719	Sequence 15719, A
14	567	23.2	2713	20	US-10-425-115-75763	Sequence 75763, A
15	520.2	21.2	536	18	US-10-424-599-32308	Sequence 32308, A
16	499.6	20.4	680	18	US-10-424-599-32307	Sequence 32307, A
17	489	20.0	2336	18	US-10-424-599-52175	Sequence 52175, A
18	482.2	19.7	3201	20	US-10-425-115-44188	Sequence 44188, A
19	477.4	19.5	1992	18	US-10-425-114-13333	Sequence 13333, A
20	465.2	19.0	2242	20	US-10-425-115-40196	Sequence 40196, A
21	433.8	17.7	585	18	US-10-424-599-108738	Sequence 108738, A
22	414.6	16.9	2251	19	US-10-437-963-72062	Sequence 72062, A
23	401.4	16.4	1980	17	US-10-260-238-1066	Sequence 1066, Ap
24	396.8	16.2	1809	18	US-10-424-599-83137	Sequence 83137, A
25	389.6	15.9	2506	18	US-10-424-599-74393	Sequence 74393, A
26	372.4	15.2	1303	18	US-10-424-599-139734	Sequence 139734, A
27	361.8	14.8	1247	18	US-10-424-599-130951	Sequence 130951, A
28	361.2	14.7	2156	20	US-10-424-599-60281	Sequence 60281, A
29	359.6	14.7	2156	20	US-10-739-930-3837	Sequence 3837, Ap
30	348.8	14.2	1551	18	US-10-424-599-54350	Sequence 54350, A
31	339.8	13.9	674	19	US-10-767-701-701	Sequence 701, App
32	329.4	13.5	1115	19	US-10-767-701-9435	Sequence 9435, Ap
33	323.8	13.2	3252	19	US-10-437-963-78357	Sequence 78357, A
34	315.4	12.9	1019	19	US-10-767-701-11661	Sequence 11661, A
35	311.6	12.7	1069	20	US-10-425-115-3619	Sequence 3619, Ap
36	298.6	12.2	2913	19	US-10-437-963-76453	Sequence 76453, A
37	298.2	12.2	1174	20	US-10-425-115-75765	Sequence 75765, A
38	293.2	12.0	2346	20	US-10-425-115-2539	Sequence 22539, A
39	291.4	11.9	1632	18	US-10-424-599-74389	Sequence 74389, A
40	285.6	11.7	565	18	US-10-424-599-62889	Sequence 62889, A
41	271.4	11.1	788	20	US-10-425-115-75767	Sequence 75767, A
42	258.8	10.6	1883	18	US-10-424-599-103813	Sequence 103813, A
43	255	10.4	459	18	US-10-424-599-120407	Sequence 120407, A
44	249.8	10.2	735	18	US-10-424-599-35721	Sequence 35721, A
45	249.4	10.2	841	18	US-10-424-599-135930	Sequence 135930, A

ALIGNMENTS

RESULT 1

US-10-424-599-131116
; Sequence 131116, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131116
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89404C.1
US-10-424-599-131116

Query Match 62.1%; Score 1521.6; DB 18; Length 2008;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 84; Indels 2; Gaps 1;
Qy 780 CCCAAACTCATGAGTGGAGTGGAAAGTCTGTTAGGATGTTCTTTCATTTCCTCC 839
Db 1 CCCAAACGCATGAGTGGAGTGGAAAGCGCTGTTGGTGTGTTCTGTTTCCTCC 60

```
QY 840 TCCTTAGCACAGATACCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTACGCAATGG 899
Db |||||
QY 900 GGCATTGACGTCGTTATATTGGGAAGTCTCTTGGTTATTTCACTCAGCGCGAGAAGC 959
Db |||||
QY 960 ACCGTGTTGAA--GTGATAGGAACTGAAGAGGGTTTGAATCCACCACATCACTCAAAA 1017
Db |||||
QY 1018 TCTGGTATTTGTGCGCTTTACATGACTACAGCTGTCAAACTCGGCATTTGCGTTGGCAT 1077
Db |||||
QY 1078 CATATCACTTCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATATAAAATTAACAA 1137
Db |||||
QY 1138 TATTGATGGCAACAAGAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTTAC 1197
Db |||||
QY 1198 CTCTTGCTACCTCACAAACAGGACCATTTTCGCGTTTCGGCTGTGAACTATAACGCTGGATG 1257
Db |||||
QY 1258 CAAGACAGCAGCTCCAACTTATATGCTCACTTGCAGTAATGTTGACATTTGTTATTCCT 1317
Db |||||
QY 1318 GACACCTCTGTCATTTACACTCCCTCGTGTGCTATCAGCTATTATCGTATCTGCAAT 1377
Db |||||
QY 1437 GCTTGGACTCATAGATTATGAAGCAGCCATCCACTATTTTAAGGTTTCACAAATTTGACTTT 1437
Db |||||
QY 1438 TGTGGTGTGCATGAGTGATACATTGGCGTGGTCTTTGGCAGTGGTTGAAATTTGGCTTAGT 1497
Db |||||
QY 1498 CATAGCTATTGTAATCTGTACTTCGGGTACTTCTATTATTATTCGAAGGCCAAGGACATT 1557
Db |||||
QY 1558 CGTTTTGGGCAACATTCCAAATTTCTGTGATATACCGAAATGTTTGAGCACTATCAAAATGC 1617
Db |||||
QY 1618 AAAACATGTCCTGGATGCTAATCTAGAGATTGATGACCAATTTACTTTGCCAATGC 1677
Db |||||
QY 1678 CAGCTATTTAAGAGAAGGATCAAGGTGATTGATGAAGAAGAAGAAATTAAGC 1737
Db |||||
QY 1738 TACAGGGGAGACTAGTTTGCAGTATGTTATTAATTTGATATGAGTGTCTTTGGAAACATTGA 1797
Db |||||
QY 1798 TACAAGTGGNAATAGTATGCTTCAAGAGGTGAAGAATTAACAGAGAGAGAGAGCTACA 1857
Db |||||
QY 1858 GCTTGTGTTTGGTCAATCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCA 1917
Db |||||
QY 1917 GCTTGTGTTTGGTCAATCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCA 1917
Db |||||
QY 1918 AAATCAATTTAGGAGAAAAATGGATCTATCTGACTGTTGAAGAGCGGTTGGAGCATGCAA 1977
Db |||||
QY 1978 CTTCAATCTACGTGCAAGCAAAACGAAACCAAGAAAGATGAAAACAGAGGGTTGGAACAA 2037
Db |||||
QY 2038 TGTGTGACTGAGTCAATATGCCAAAGAGTATTCTAAATAACTCAAAAGCTTATTCGTTTT 2097
Db |||||
QY 2098 CGTCTTAGTAATGTTTACCATACTCAATGTGTGGCATGAGAATTTCTGAAATCAAGCCGAAGA 2157
Db |||||
QY 2158 AGTTTTAAAGGCATAGGAAAAATGAAAGATGCAAGGGTCTTTCTAATTTCTCAACTCTGCAT 2217
Db |||||
QY 2218 CTTAGTTAGAGAAAAATCTCTATGTATAGCTGTTGAAATAACTCTTTACGTTATCATGC 2277
Db |||||
QY 2278 TTGATAATAATATTCAGAGAAAAATGTCAGCAACACTCTCAGACACACTCTTTTGAACAC 2337
Db |||||
QY 2338 ATGTAAGAGGTAAAGAGTGTGTTGCTAGCACTCTCCATATTCAATTTGTAAGTAAT 2397
Db |||||
QY 2398 GCATGAGAAATTTAAAAATCCTTTTGAAAAAATAAAAAAATAAAAAA 2447
Db |||||
QY 2447 GCATGAGAAATTTAAAAATCCTTTTGAAAAAATAAAAAAATAAAAAA 2447
Db |||||
```

RESULT 2

```
US-09-938-842A-1502
; Sequence 1502, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Kun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1502
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1502
```

Query Match 41.6%; Score 1018.2; DB 9; Length 1977;

Best Local Similarity 72.3%; Pred. No. 4e-263;

Matches 1364; Conservative 0; Mismatches 518; Indels 5; Gaps 3;

QY 125 CACCAGGTGCGAGTCTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG 184

Db |||||

QY 58 CACAGGTGGAGCTCCACAACTCAACCGTCTTGAAGTCACTTCAGTACTCAGTGAAG 117

Db |||||

QY 185 GAGACTTTCTTCCCTGTGATGACCCCTTTGAGGCGAGTTTCAAGAACAAAGCCAGCTTCCCAAGAA 244

Db |||||

118	Db	 GAAACTCTGTTTCCAGACGACCCCTTTTAGACAATTTAAGAACCAAAAATGCATCAAGAAA	177
245	Qy	TTCAATGCTGGCCCTTCAGTTCTTCTCCCAATTTTGAATGGGCTCCCAATACACCTTT	304
178	Db	TTTGTGTTAGGCTCAAAATCTTCTCCGATTTTCGATGGGACACACGCTACATCTC	237
305	Qy	CAGTTCTGAAAGCTGACCTCATAGCTGGCATCAACATCGCTAGCTTGGGCATTCCTCAG	364
238	Db	AAGTTCTTCAAAATCAGATCTCATCGCGGAATACCAATCGCTAGCTCGCATCCCTCAG	297
365	Qy	GGCATCAGTTATGCAAGCTCGCAACCTCCCTCCAATTTCTTGCATATATTTCGAGCTTT	424
298	Db	GGCATCAGTTACGCGCAAACTTGTAACTTGCCCCAAATTTCTGGCCCTTTATTTCGAGTTT	357
425	Qy	ATACCAACATTTGATTTATGCGATGATCGGTAGCTCGAGGGATTTGGCAGTGGGAGCTGTG	484
358	Db	GTACCGCCATTGTTATACGCGGTCTAGGGAGTTCAAGGGACTTAGCGTGGGACGCTT	417
485	Qy	CGGTTTGGATCGCTTCTTGATGGGTTTCCGATGTGTAGTAAATGCGGTGATCCCAATGAAGAC	544
418	Db	CGGTTTCGCTCTTGTGTGACAGGTGCGATGCTGAGCAAGAAAGTTGATGCTGAGAAAAGT	477
545	Qy	CCAAAGCTTTTACCTCCACCTGGCTTTTACAGCTACATTTATTTGCTGGTGTT-TTTCAGGC	603
478	Db	CCTAAGCTTTTACCCTTTACCTTGCTTTTCCCGCACTTTTTCGCGCGCTTCGAAAGCC	537
604	Qy	TGCCTTGGGCTCTGTTTAGTTAGGTTGGGGTTGATCGTGGAATTTTCTGTCAATGCAACATAAT	663
538	Db	TCCTTTGGAATTTTACGGTTTAGGGTTCTAGTGGACTTTCTATCGCATGCAAGATAGT	597
664	Qy	AGGTTTCATGGGAGGAGCAGCCAGGTGGTGTGTCTGCAGCAACTAAAAATCGATTTCTTGG	723
598	Db	AGGATTCATGGGAGGAGCAGCGAGGTGGTGTGAGTCTGCAACAGCTTAAGGGTATTTTCGG	657
724	Qy	CTTTGACATTTTACCCATGGAGCTGATATCATATCAGTGATGGCTCTGTTTTCACCCA	783
658	Db	ACTTAAACATTTTACAGACTCTACCGATGTTATCTCTGTCTGTCAGCGTTCCGTTTCTCCCA	717
784	Qy	AACCTCATGAGTGGAGTGGGAAAGTCTGTGTTAGGATGTCTCTCATTTTCTTCCCTCT	843
718	Db	AACCTCAGTGGAGATGGGAAAGTGGCTTCTTGGATGGTTCATTTCTTTCTTCTCT	777
844	Qy	TAGCACAAGATACCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGCC	903
778	Db	CTCCACCAGATATTTTGAGCATCAAGAAACCAAAATTTCTTTTGGGTGGCGGATGGCTCC	837
904	Qy	ATTGACCTCGTTATATTTGGGAAGTCTCTGGTTTATTTCTCATCGCCGAGAAAGCAGG	963
838	Db	TTTGACCTCAGTGATTTCTGGAAGTCTCTTGGTTTACTTCACTCAGCTCAGAGACATGG	897
964	Qy	TGTTTGAAGTGTAGGAGAACTGGAAGAAGGTTTGAATCCACCACTCAACAATCTGGT	1023
898	Db	TGTTTCAAGTGAATGGGACCTTGAAGAAGGTTGAATCCACTCTCCGGTTCTGATCTCAT	957
1024	Qy	ATTTGTCTCGCTTATCACTACAGCTGTCAAACTGGCATTTGCTGTGGCATCATATC	1083
958	Db	CTTTTACTTCCCTTACATGTCCACAGCTGTCAAACTGGCTCATCTACCTGGCATTTGC	1017
1084	Qy	ACTTGGCGAAGGAATAGCAGTAGGAAGAGCTTTTGAATGTATAAAATTTACAATATGA	1143
1018	Db	TCTCGC-GAAGGAGTAGCAGTGGGGAGGATTTTTCGATGTTTCAAGAACTACAACATAGA	1076
1144	Qy	TGGCAACAAAGAGATGATAGCTATTTGGGACCATGAAAGTGTGGTTCTTTCACTCTTG	1203
1077	Db	CGGAAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCTTTCATCTTG	1136
1204	Qy	CTACTCACAACAGGACATTTTTCGGCTGTGAACTATATAACGCTGGATGCAAGAC	1263
1137	Db	TTACTCACAACCGGACCAATTTTCAAGTGGCAGTGAACCTACAAACGGGGTTCAGAGAC	1196
1264	Qy	AGCAGCTTCCAAACANTATATGTCATTTGCAGTGAATGTTGACATGTTGTTTCCTGACACC	1323

Db	1197	CGCAATGTCACACATAGTGTATGGCGATTATGTTCCACACTCTCTTCTCTCACACC	1255
Qy	1324	CTTGTTCCATTACACTCCCTCGTGGTGTCTATCAGCTATTATCGTATCTGCAATGCTTGG	1383
Db	1257	GCYTTTTCACTACACACCACTCGTGTGCTCTCTGCCATCATCATATCCGCAATGCTCGG	1316
Qy	1384	ACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGTTTGACAAATTTGACTTTTGCGT	1443
Db	1317	ACTCATTTGACTATCAAGCTGCCATCCATCTCTCGAAAGTTTGACAAAGTTCCGACTTCTCTCGT	1376
Qy	1444	GTGCATGAGTGCATACATTCGCGTGGTCTTTTGGCAGTGTGAAATTTGGCTTTAGTCTATGAC	1503
Db	1377	CTGCATGAGCGCCTACGCTGGGGTCGATTTCGGCAGGTAGAGATTGGACTCGTCGTGAC	1436
Qy	1504	TATTGTAATATCTGTACTTCGCGGTACTTCTATTTATTTGCAAGGCCAAGGACATTCGTTTTT	1563
Db	1437	GGTGGCGATATCTATAGCGAGGTTGTTGCTGTTTGTGTCGAGGCCAAAACTCGGTTGAA	1496
Qy	1564	GGGCAACATTTCCAAATTCCTGTGATATACCGAAATGTTTGAGCAGCTATCAAAATGCAAAACA	1623
Db	1497	GGGAAACATATCCAAACAGCATGATCTATAGGACACCTGAGCAGGTACCCATCATCAAGAAC	1556
Qy	1624	TGTTCTCTGGAAATGCTAATTCCTAGAGATTGATGCACCAATTTTACTTTTGCCCAATGCCAGCTA	1683
Db	1557	CGTTCCTCGGTATCTCATCTTTGGAGATTGATGCTCCCATCTACTTTGCTAATGCCAGTTA	1616
Qy	1684	TTTTAAGAGAAGGATCAACAGGTGGATTGATGAAGAGAAGAAAGAAATTTAAAGCTACAGG	1743
Db	1617	CTTGGTGAAGAATCATAGGTGGATTGATGAAGAGAGAGAGATTAAACAATCAGG	1676
Qy	1744	GGAGACTAGTTTGCAGTATGTTTAAATTTGATATGATGCTGTTGGAAACATTGATACAAG	1803
Db	1677	AGAGAGCAGCTTACAATATATTTACTTCGATATGTCAGCTGTTGGTAAATATCGACACAAG	1736
Qy	1804	TGGAAATGATGCTCTGAAGAGGTGAAGAGATTACAGAGAGAAGAGAGCTACAGCTTGT	1863
Db	1737	CGGTATTAGCATGATGTTGGAAATTAAGAAGTTCATTGACAGGAGAGCGTTTAAAGTTGGT	1796
Qy	1864	TTTGGTCAATCTCTGAAGTGAAGTGAAGAAACTGAACAAAATCGAAGT---TCCAAAA	1920
Db	1797	ATTGTCAATTCCAAAGGAGAGGTGCTGAAGAAATTAACAGATCCAAATTCATCGGTGA	1856
Qy	1921	TCATTTAGGGAAGAAATGGAATCTATCTGACTGTTGAAGAGCGCGTTGGAGCATGCAACTT	1980
Db	1857	TCATTTGGCCAAAGAGTGGATGTTCTTAAACGTTAGGAGACAGTGGAGGCTTTGTAGCTA	1916
Qy	1981	CAATCTACGTGCGAAGCAAAACGAACC	2007
Db	1917	CATGCTTTCACACGTTTAAAAACCGAACC	1943
RESULT 3			
US-09-938-842A-1502			
; Sequence 1502, Application US/09938842A			
; Publication No. US20040009476A9			
; GENERAL INFORMATION:			
; APPLICANT: Harper, Jeff			
; APPLICANT: Kreps, Joel			
; APPLICANT: Wang, Xun			
; APPLICANT: Zhu, Tong			
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING			
; TITLE OF INVENTION: SAME, AND METHODS OF USE			
; FILE REFERENCE: SCRIPI300-3			
; CURRENT APPLICATION NUMBER: US/09/938,842A			
; CURRENT FILING DATE: 2001-08-24			
; PRIOR APPLICATION NUMBER: US 60/227,866			
; PRIOR FILING DATE: 2000-08-24			
; PRIOR APPLICATION NUMBER: US 60/264,647			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/300,111			
; PRIOR FILING DATE: 2001-06-22			
; NUMBER OF SEQ ID NOS: 5379			
; SEQ ID NO 1502			

RESULT 3

RES-001.3
 US-09-938-842A-1502
 ; Sequence 1502, Application US/09938842A
 ; Publication No. US20040009476A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRP1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1502

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

RESULT 4

US-10-437-963-38184
; Sequence 38184, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua


```

RESULT 6
US-10-425-115-50601
; Sequence 50601, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 50601
; LENGTH: 2727
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_146143C.1
US-10-425-115-50601

Query Match          27.7%; Score 677.4; DB 20; Length 2727;
Best Local Similarity 60.8%; Pred. No. 3.6e-171;
Matches 1123; Conservative 0; Mismatches 721; Indels 3; Gaps 1;

OV      122  GTGCACCAAGTCGAGGTTTCACCCGCCACAGCGCGTTTTTCAAGTCTCTAAAGTACTCTTTG 181

```

Db 300 GGGTACAAAGTCGGCGTCCCAACCGGAGAGAAACCTCTCTCGCCGAGATCTCTGACGCGGTG 359
Qy 182 AAGGAGAGCTTTCTTCCCTGATGACCCCTTTGAGGAGCTTCAAGAAACAAGCCAGCTTCCAAAG 241
Db 360 AAGGAGAGCTTTCTCGCGAGACCCGCTGGGAGTACAGGACCAAGCCAGCTCCAG 419
Qy 242 AAGTTCATGCTTGGCTTCAGTTCCTTCTCCCAATTTTGAATGGGTCCCAATACAC 301
Db 420 AAGATCTGGCTCGGCTCGAGCAATCTTCCGCTGCTCGAAATGGAGCAGGCATTAATCC 479
Qy 302 TTTTCAGTCTTGAAGCTGACCTCATAGCTGGCATCACCTCGTAGCTTGGCATTCCT 361
Db 480 CTGGCAAGTTCAAGGCTGACTTCTCGCTGGGCTCACCATTTGCCAGCTCTGCATACCC 539
Qy 362 CAGGCGATCAGTTATGCAAGCTGCCAACCTCCCTCCCAATTTCTTGGACTATATTCGAGC 421
Db 540 CAGGACATCGGTATTCTAAGCTTGCTAACTTCCAGCAGAGGTGGGACTATACAGTAGC 599
Qy 422 TTTATACCAATGATTTATGCGATGATGGTATGCTCGAGGGATTTGGCAGTGGGACT 481
Db 600 TTTGTCGCGCTCTGATATAGCAGTGTATGGCAGCTCCAGGACATAGCCATCGGTCCA 659
Qy 482 GTGCGGTGATCGCTTCTGATGGGTTCGATGTTGAGTAAATCCGTTGATCCCAATGAA 541
Db 660 GTGCGGTGATCGCTTCTGATGGGTTCGATGTTGAGTAAATCCGTTGATCCCAATGAA 719
Qy 542 GACCCAAAGCTTTACCTCCACCTGGCTTTCACAGCTACATTAATTTGCTGGTGTTCAG 601
Db 720 CACCCGCTCGAGTACAGGCGCTAGCCCTTACAGCAACCTTTTTCGAGGCTCACTCAG 779
Qy 602 GTCGCTTGGGTCTGTTTATGCTGGGTGATGCTGATGTTTCTGTCATGCAACCAATA 661
Db 780 GCAGCCCTCGGATCTTTCAGGCTAGGCTTTATCATAGATCTTGTCTCATGCTGCCATT 839
Qy 662 ATAGGTTTCATGGAGGAGCAGCACCGTGGTGTCTGAGCAACCAATAAATCGATCTTT 721
Db 840 GTCGGAATTCATGGCGCGCTGCCATCACCATTGCGCTTCAGCAGCTGAAAGGATTCCTT 899
Qy 722 GGCCTTGAGCATTTACCCCATGAGCTGATATCATATCAGTGAATGCGCTCTGTTTTCAAC 781
Db 900 GGAATTCGAATTTACCAAGAAATCCGATATGATCTGTTATGAATCAGTTGGGA 959
Qy 782 CAAACT---CATGAGTGGAGTGGAAAGTGTGTTGATGATGTTCTTCAATTTCTTC 838
Db 960 AATGTTCAACATGGTGGAAATGGCAGCAATACTGATAGGAGCGACCTTCTGGCATTC 1019
Qy 839 CTCCTTAGCACAATGATCTTACGAAACCAAGCAAGGTTTTTTGGGTGTCAGCAATG 898
Db 1020 CTTCTGTTGCCAAGTACATTTGGGAAAGGAATAAAAGGCTCTCTGGGTGTCGCAATC 1079
Qy 899 GCGCATTTGACGTCGGTTATTTGGGAAGTCTCTTGGTTTATTTTCACTCACGCGGAGAG 958
Db 1080 GCACCTCTCACTTCGGTGATCATATCCACATTTTGTGTACATCACTCGTGAGATAAG 1139
Qy 959 CAGGTTGTTGAAGTATAGAGAACTGAAGAGGTTTGAATCCACCATCACTCACAAAT 1018
Db 1140 CATGGCTTGCATTTGTCAAGAAATGAAGAAAGGCAATCAACCCACCTTCAGCTAGCTA 1199
Qy 1019 CTGGTATTTGTGTCGCTTACATGACTACAGCTGTCAAACTGGCATTTGCTGGCATC 1078
Db 1200 ATATACTTCAACCGGCCCATACTTGGCAGCAGGATTCAAATTTGGGATGAGTGGGAATG 1259
Qy 1079 ATATCACTTCGGGAGGAATAGCAGTATAGGAAGAGCTTTGCAATGTATAAAATTAACAAT 1138
Db 1260 ATAGGCTTAACGGAGCGATTTGCAATTTGGAAGAACATTTTCAGCCCTCAAGGATACCGG 1319
Qy 1139 ATTGATGGCAACAAGAGATGATAGCTATTGGGACCAATGAACGATGTTCTTTCACC 1198
Db 1320 ATAGATGGGAACAAGAAATGATGGCTTAGGGACCAATGAATTTGTTGGTTCACTGACT 1379
Qy 1199 TCTTGCTACCTCACACAGGACCAATTTTCGGGTTCGGCTGTGAATTAACGCTGGATGC 1258

Db 1380 TCTTGCTACTAGCCACAGGTTCTTTCTCGCGTCAAGCTTAATTTACATGCGCTGGCTGC 1439
Qy 1259 AAGACAGCAGCTTTCCAACTAATAATGTCATCTTGCAGTAAATGTTGACATTTGTTCTCTG 1318
Db 1440 ABAACAGCAGTGTCAACCGTTGTTGTCATTAATGTCGTAATGCTTACTTTGCTACTGATC 1499
Qy 1319 ACACCTTGTTCATTTACACTCCCTCGGTGGTGTATCAGCTATTATCTGATCTGCAATG 1378
Db 1500 ACCCATTTGTTCAAGTACACTCCAAATGCCATCTTTCTCTCGATCATATCAGCAGTG 1559
Qy 1379 CTTGGACTCATAGATATGAGCAGCAGCATCTCTATTTAAAGGTTGCAAAATTTGACTTTT 1438
Db 1560 CTTGGTTTAAATGACTATGAAATCGGCTTACCTTATCTGAAAGTTGACAACTGGACTTT 1619
Qy 1439 GTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1498
Db 1620 CTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1679
Qy 1499 ATAGCTATTGTAATATCTGATCTTGGGTACTTCTATTTATTTGCAAGGCCAAGACATTC 1558
Db 1680 ATTGCGGTGCAATATCTCTTGTCTAAATTTCTTCCAAAGTAAACGCGCCCAAGACAGTT 1739
Qy 1559 GTTTTCGGCAACATTTCCAAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATGCA 1618
Db 1740 TTAATTTGGAACCTTCCAGCAACATATATACGGAATGTAGAACAGTATCTCTGATGCT 1799
Qy 1619 AAAACATGTTCTTGGAAATGCTTAATTTAGAGATTTGATGCACTTCTTTCGCAATGCC 1678
Db 1800 ACCAAGTTCCAGGGTGTGATTTGTTAGAGTGAATCAGCTATATCTTTCACAACTCT 1859
Qy 1679 AGCTATTTAAGAGAAAGGATCAAGGTGATGATGATGATGATGATGATGATGATGATGAT 1738
Db 1860 AACTATGTTTAAAGAGAGAAATCTTGGAGTGGCTAAGAGATGAGGAGGAGCAACACAGGAC 1919
Qy 1739 ACAGGGGAGCTAGTTTTCAGTATGTTTAAATTTGATGATGATGATGATGATGATGATGAT 1798
Db 1920 CAGAAGTTTAAACAAACTGATTTCTAATTTGATGATGATGATGATGATGATGATGATGAT 1979
Qy 1799 ACAAGTGGAAATAGTATGCTTGAAGAGTGAAGAGATTTACAGAGAGAGAGAGTACAG 1858
Db 1980 ACAAGTGGAAATCCATGCTTTTGGAGGAGTTGGCGAAAGCTTTTGAAGAAATTCAG 2039
Qy 1859 CTTGTTTTGGTCTAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1918
Db 2040 CTGGTTCTTACCAATCCCGGCGCGGTGATCCAGAAGCTCGCTCAGCGAAATTCAGC 2099
Qy 1919 AATCATTTAGGGAAGAAATGATCTATCTGATCTGATCTGATCTGATCTGATCTGATCT 1965
Db 2100 GACATGATTGGTGAAGAACAACTATTTCTCCTCAGCGTGGCGACGCTGT 2146

RESULT 7

US-10-437-963-77012
; Sequence 77012, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77012
; LENGTH: 4390
; TYPE: DNA

; ORGANISM: Oryza sativa									
; FEATURE:									
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76952C.1									
US-10-437-963-77012									
Query Match 27.4%; Score 670.4; DB 19; Length 4390;									
Best Local Similarity 60.3%; Pred. No. 3,7e-169;									
Matches 1109; Conservative 0; Mismatches 731; Indels 0; Gaps 0;									
QY	126	ACCAAGTCGAGGTTCCACCGCCACACGCGTTTTCAAGTCTCTAAAGTACTCTTTCAAGG	185						
DB	504	ACAATGTCGCGCGCGCCGCAAGAAGAACTCTCTGCGGAGTTTCGCCGACCGTGAAGG	563						
QY	186	AGACTTTCTTCCCTGTATGACCCCTTTGAGGCAGTTTCAAGAACAAAGCCAGCTTCCAAAGAT	245						
DB	564	AGAGGTTCTTCGACGAGCCGATGCGGGGTACAAGGACACGCGAGGTCCAGGAAGC	623						
QY	246	TCATGCTTGGCCCTTCAGTTCTTCTTCCCAATTTTGAATGGGCTCCCAATACACCTTTC	305						
DB	624	TATGGCTCGCTTGCAGCAGCTCTTCCGGTGTTCGAATGGGCGACACAATACACCCCTCG	683						
QY	306	AGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAGG	365						
DB	684	CCAAGTTCAAGGGCGACCTCAITGGCGGCTCACCCTTGCACGCCCTCGTCATACCTCAGG	743						
QY	366	GCATCAGTTATGCAAGCTCGCCAACTCCCTCCAATTTCTTGAGCTATATTCGAGCTTTA	425						
DB	744	ACATCGGTACGGAGCTTGCTTAACCTGCCACACAGATTTGGCTGCGACAGTAGCTTCG	803						
QY	426	TACACCATTTGATTTATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAGCTGTGG	485						
DB	804	TCCCGCATTTGATATACGCTCTGATGGGCACTCAAGGGAGCTAGCAATGGGTCCAGTGG	863						
QY	486	CGGTTGATCGCTTCTGATGGGTTGATGTTGAGTAAATCCGCTTGATCCCAATGAAGACC	545						
DB	864	CCGTCACTCACTGCTTGGTACTCTCTCCAGGAGGAGATTGACTCAAGAAGAACC	923						
QY	546	CAAAAGCTTTACCTCCACCTGGCTTTCAAGCTACATTTATTTGCTGGTGTTTTTCAGGCTG	605						
DB	924	CGCTAGATTACAGCGGCTCGCCTTACAGCGACCTTCTTTGAGAGTCAACAGCGCG	983						
QY	606	CCTTGGGTCGTGTTAGTTGGGTTGATCGTGAATTTCTGTCAATGCAATCAATATAG	665						
DB	984	CGCTGGGTTTCTCAGGCTAGGGTTTCATCATAGCTTCTGTCTCATGTGCCATCATCG	1043						
QY	666	GGTTCAATGGGAGGAGCAGCAGCGTGTGTCTCGACGAACCTAAATCGATTCTTGGCC	725						
DB	1044	GATTCATGGCCGGCGGCCATCACCATTGCTCTTCAGCAGCTTAAAGGCTTCCITGGAA	1103						
QY	726	TTGAGCATTTACCCATGGAGCTGATATCATATCAGTGATGCGCTCTGTTTTCCACCCAAA	785						
DB	1104	TTGCAAACTTCCAAAGAGACTGACATCATCTCGGTCTATGAATCAGTCTGGGAAATG	1163						
QY	786	CTCATGATGGAGGTGGGAAAGTGTGTGTAGGATGTGTCTCATTTTCTTCCCTCTTA	845						
DB	1164	TTCAACATGGGGCAATGGAACTGGCAGACAAATATTGATCGGAGCATCATTTTGGCATTC	1223						
QY	846	GCAACAGATATTCAGCAAAAACGACCAAGGTTTTTTGGGTGTGAGCAATGGCGCCAT	905						
DB	1224	CTCCTGGTTGCCAAGGCAAGAGAAACAAGAAGCTCTTCTGGGTCCCAAGCAATGGACCAC	1283						
QY	906	TGACGTCGGTTATATTGGGAAGTCTCTTGGTTTTATTTCATCTACGCGCGAGAACGCGTG	965						
DB	1284	TCATTTCCGGTGATCATTTCAACTTTGTTGCTGTACATCATCTGTGTGACAAACAGGCG	1343						
QY	966	TTGAAGTGATAGAGAACTGAAGAAGGTTTGAATCCACATCACCTCAAAATCTGGTAT	1025						
DB	1344	TCGCAATCGTAAAAAAGCTCAAGAAAGGCAATCAATCCACCTTCAGCTAGCTGATATTT	1403						
QY	1026	TTGTGTGCTTACATGACTTACAGCTGTCAAACTGSCATTTGCTTGGCATCATATCAC	1085						
DB	1404	TCACTGGCCCATACTTGCTCAAGGATTCAAAATTGAGTAGTAGCTGGAATGATGAAGCC	1463						

QY	1086	TTTCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTACAATATTGATG	1145
DB	1464	TTACCGAAGCGATTGCAGTTGGAAGAACATTTTGTGCGATTGAACGATTACCAGATAGATG	1523
QY	1146	GCAACAAAGAGATAGTATTTGGGACCATGAACTAGTGTGTTCTTTTCACTCTTGTCT	1205
DB	1524	GGAACAAAGAAATGTTGGCTCTAGGAACCATGAATGTGTTGTTCAATGAGCTCTTGTCT	1583
QY	1206	ACCTCACAACGAGCCATTTTCGGCTTGAACTATACGCTGATGCAAGACAG	1265
DB	1584	ATATAGCCACAGTGGTTTTTGACGATCAGAGTCAATTTGCTGCTGGAGGTAAACAC	1643
QY	1266	CAGCTTCCAACTATTAATGTCTATTCAGTAAATGTGACATTTGTTATTTCTTGACACCT	1325
DB	1644	CAATGTCCAAATTTGTTATGTCACTGTAGTATTTGCTTGCACTCTCTGTGGATCACTCCAT	1703
QY	1326	TGTTCCATTTACACTCCCTGGTGGTCTATCAGCTATTATTCGTATCTGCAATGCTTTGGAC	1385
DB	1704	TGTTCAAAGTACACGCGCCAAATGCCCAATTTCTTCCATCATATATCAGCAGTGTGGCC	1763
QY	1386	TCATAGATTTATGAAGCAGCCATCCATCTATTTAAAGTTGACAAATTTTGACTTTTGGTCT	1445
DB	1764	TATTTGACTTTGAATCAGCTTACCTTATCTGGAAGGTTGATAAGTTGGACTTCAAGCGT	1823
QY	1446	GCATGAGTCATACATTTGGCGTGGTCTTTTGGCAGTGTGAAATTTGGCTTTAGTCATAGCTA	1505
DB	1824	GCTTGGGGCATTCCTTGGAGTAATATTTTCATCTGTGGAGTATGGCTTCTCATTTGCGG	1883
QY	1506	TTGTAATATCTGTACTTTCGGGTACTTCTATTATTATTTGCAAGGCCAAGGACATTCGTTTGG	1565
DB	1884	TTGTAATATCACTAATCAAAAGTTCTGCTCCATGTAAACACGCGCCAAAGGACAGCTTTACTTG	1943
QY	1566	GCAACATTTCCAAATTTCTGTGATATACCGAAATGTTGAGCAGCTATCAAAATGCAAAACATG	1625
DB	1944	GCAACCTTCCAAAGACGATTATCTATAGAAATGTTGAACATATCCAGAGCTACCAAG	2003
QY	1626	TTCTCGAATGCTAAATTTCTAGAGATTGATGCAACCAATTTTACTTTGCAATGCCAGCTATT	1685
DB	2004	TGCCAGGATGCTAATTTGTAAGAGTGGACTCAGCAATATACITTCACAACTCCAATTATG	2063
QY	1686	TAAGAGAAAGGATCACAAGGTGGATTGATGAAGAGAGAGAAATTAAGCTACAGGGG	1745
DB	2064	TTAAAGAAAGAAATGCTGAGATGGCTGAGAGATGAGAGAGAAACATCAAAAGGAACAGAAT	2123
QY	1746	AGACTAGTTTGCAGTATGTTATAATTGATGATGAGTGTGTTGGAAACATTTGATACAAAGT	1805
DB	2124	TACCAAAATTTGAGTTTCTGATTTGACCTATCTCTGTAATGATATTGACACAGTG	2183
QY	1806	GAATAAGTATGCTTGAAGAGGTGAAGAGATTTACAGAGAGAGAGAGCTTACAGCTTGT	1865
DB	2184	GAATCCATGCCCTTCAAAGAGTTGTTGAGGACACTTTGAAAAGCGCCAGATTACGCTGATTT	2243
QY	1866	TGGTCAATCTCTAAGTGAAGTGAAGAACTGACAAATGCAAAATCGAAGTTCCAAATCATTT	1925
DB	2244	TGCGCAATCTCGGGCGGCTGTGATCCAAAAGCTCCGGTCAGCAAAATTCACAGAGCTCA	2303
QY	1926	TAGGGAAGAAATGAGTCTATCTGACTGTTTGAAGAGGCGCT	1965
DB	2304	TTGGTGAAGAAAGATATGCCCTGCACAGTTGGTGACGCCGT	2343

RESULT 8
US-10-424-599-65103
; Sequence 65103, Application US/10424599
; Publication No. US200400031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599									
; CURRENT FILING DATE: 2005-04-28									
; NUMBER OF SEQ ID NOS: 285684									
; SEQ ID NO 65103									
; LENGTH: 2366									
; TYPE: DNA									
; ORGANISM: Glycine max									
; FEATURE:									
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29800C.1									
US-10-424-599-65103									
Query Match 26.9%; Score 658.6; DB 18; Length 2366;									
Best Local Similarity 60.5%; Pred. No. 3.9e-166;									
Matches 1120; Conservative 0; Mismatches 724; Indels 7; Gaps 2;									
Qy	77	AGTGTAGATTATGAGTACCCCTTTGGGCATGAACAACATTTTGAGAGAGTGCAACCAAGTCGAG	136						
Db	132	AATGCATGTACTATGCATTTCTCATTCATGATAGATGTCATGGAAGTGCAACCAAGTTGTG	191						
Qy	137	GTTCACCGCCACAGCCGTTTTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACTTTCTTTC	196						
Db	192	CCACCACCTCACAGAGCACCCCTTCAGAAACTCAAGGGTAGGCTCAAGGAAACTTTCTTTC	251						
Qy	197	CCTGATGACCTTTGAGGCGAGTTCAAGAACAGCCAGCTTCCAAAGATTCTGCTTGGC	256						
Db	252	CCTGATGATCCTCTGCGCCCAATTCAGGGACAACCTCTTAAGAGAAAACCTGATCCCTTGG	311						
Qy	257	CTTCAGTCTTCTTCCCAATTTTGAATGGGCTCCCAATACACCTTTTCAGTTCTTTGAA	316						
Db	312	GCTCAATATGTGTCCCTATTTCTCAATGGGCTCTAAGTATTAATCTCAAACTCTTCAA	371						
Qy	317	GCTGACCTCATAGTGGCATCACATCGTAGCTTGGCCATTCCTCAGGGCATCAGTTAT	376						
Db	372	TCTGACCTTGTCTGCGCTCACTATTGCTAGCTTGACATCCCGCAGGGAATGAGTTAT	431						
Qy	377	GCCAAGCTCCCAACCTCCCTCAATTTCTGGACTATATTCGAGCTTTATACCAACATTC	436						
Db	432	GCTTAGCTTGCAGTCTTCTCCAAATTTTAGGACTTTATTTCTAGTTTGTGTCCACACAT	491						
Qy	437	ATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGGACTGTGGCGTTGGATCG	496						
Db	492	GTCTATGCTGTCTTGGAACTCAAGGACCTTGGCAGTTGGACCTGTCTTATTTGCTTCT	551						
Qy	497	CTTCTGATGGGTTCGATTTGAGTAATGCCGTGATCCCAATGAAGCAACCAAGCTTTAC	556						
Db	552	CTTTGATGGGATCCATTTGTGCATCAGGAAGTGTCTCCCAACACAGATCCAAATCTGTTT	611						
Qy	557	CTCCACCTGGCTTTCAGCTACATTTATTTGCTGGTGTGTTTTTCAGGCTGCTTGGGTCTG	616						
Db	612	CTTCAGCTAGCTTTTCACTTCAACATTTATTTGCGGCTCTTTTCAAGCTTTGCTTGGAA	671						
Qy	617	TTTAGGTTGGGGTTGATCGTGGATTTTCTGTCACTGCAACCAACCAATATAGGGTTTCATGGGA	676						
Db	672	CTAAGGCTAGGCTTCAATATGATTTTCTATCTAAGGCCATTTCTATTGGGTTTCATGGCT	731						
Qy	677	GGAGCAGCCACGGTGGTGTCTGACGAACTAAATTCGATTTCTGGCCTTTGAGCAATTC	736						
Db	732	GGAGTGTCTATTATTGTCTCACTGCAACAGCTCAAGAGCCTGCTTGGAAATCACACATTC	791						
Qy	737	ACCATGGAGCTGATCATATCAGTATGCGCTCTGTTTTTACCACCACTCATGATGG	796						
Db	792	ACTAATCAGATGGGTTCGATTTCTGTTATGACTTCTGTTTTTCACAATATACATGATGG	851						
Qy	797	AGGTGGGAAGTGTGTGTTAGGATGTCTTCAATTTCTTCTCTCTTACCAAGATAC	856						
Db	852	TCATGGCAACAATATTGATGGGATTTGCTTCTTGCTACTACTACTATTAGCAAGAC	911						
Qy	857	TTCAGCAAAAAACGACCAAGGTTTTTTGGGTGTTCAGCAATGGCGCAATTCAGCTCCGTT	916						
Db	912	GTTAGCATAGGAAACCAAACTATTCTGGGTCTCAGCTGGAGCTCTCTTATGTGTGTC	971						
Qy	917	ATATTGGGAAGTCTCTGGTTTTTATTTCACTCACGCCGAGAACCGTGTGTGAAGTATA	976						

RESULT 9
US-10-425-115-182207
; Sequence 182207, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION: Thomas J.
; APPLICANT: La Rosa, Thomas J.

1926
1925
1875
1865
1815
1805
1755
1751
1695
1691
1635
1631
1575
1571
1511
1455
1451
1395
1391
1331
1275
1271
1215
1211
1155
1151
1095
1091
1036
1031

972 ATCATCTTACCCCTCTTGGTTTTTGCATTAAGGCTCAAAATCATGGCATCAGTGCATT
977 GGAGAACTGAAGAAGGTTTTGAATCCACCATCACTCACAATCTGTATTTGTGTCGCT
1032 GGAATAATGCAACAAAGGAATAATCTCCATCATGGAATATGTGCTCTTTTCATGAAGT
1037 TACATGACTACAGCTGTCAAACTGCAATTTGCTGGCATCATATCACTTGC- GGAAG
1092 CACCTAGGCCTAGTTATGAANAACAGGGCTTATCACCGGATTTTGTCCCTAAACGAAGG
1096 AATAGCAGTAGGAAGAGCTTTTGCATGTATAAAAATTACAAATTTGATGGCAACAAAGA
1152 TATTGACGATGAGAGACATTTGCGACTCTCAAAACTACAAAGTGGATGAATAAGGA
1156 GATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTTTCACCTTTGCTACCTCAAC
1212 AATGATGGCAATTGGGTTTTATGAATGTTGTTGGCTCTTTTCTGCTATGTTTCAAC
1216 AGGACCAATTTTCCGCTTCCGCTGTGAACCTATAACGCTGGATGCAACAGCAGCTTCCAA
1272 AGGTGCTTTCTCTCGGTGAGCAGTTAAACAATGCAGCGCGCAAAACAGCTGTGTCAA
1276 CATTATAATGTCACCTTTCAGTAATGTTGACATTTGTTTATCTCTGACACCTTTGTTCAATTA
1332 TGTAGTATGCTGTGACAGTCAATGCTGACCTCTTTTCTCATGCCATTTGTTCAATA
1336 CACTCCCTGGTGGTCTATCAGTATTAATCGTATCTGCAATGCTTGGACTCATAGATTA
1392 CAGCCTAATGTCGTGCTGGGCAATCATAGTTCACAGCAGTAATTTGGCTCATCGATCT
1396 TGAAGCAGCATCCATCTATTTAAGTTGACAAATTTGACTTTGCTGGTGGCATGATGC
1452 CCCCCTGCTGTTGAACATTTGGAAGATCGACAAATTCGATTTGTTGATGATGATGC
1456 ATACATTTGCGCTGCTTTTGGCAGTGTTCAAATTTGCTTAGTCATAGCTATTGTAATATC
1512 TTTCTTAGTGTCTTTTCTCATCTGCTCAGAGGCTTGTCTTCTGCTGTTGGATATC
1516 TGTACTTCCGGTACTTCTATTTTTCAGAGGCAAGGACATTCGTTTTGGGCAACATCC
1572 AACTTTGAAGATACCTCTTCAAAATTACGAGACCGGAAACAGTGTGTTGGGAAGATACC
1576 AAATCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATGTTCTCGAAT
1632 AGGAACAGACATATATAGAAATCTTGATCAATACAGGAAGCTGTGAGAATACCTGGATT
1636 GCTAATTTAGAGATTGATGCACCAATTTTACCTTTTGCATGCCAGCTATTTTAAGAGAAAG
1692 TCTTATTTTAAAGATTGAGCTCCCATCAATTTTGCACACATCAATATCTCAACGAGAG
1696 GATCAAGGTTGGATTGATGAAGAAGAAAGAAATTTAAAGCTTACAGGGGAGACTAGTTT
1752 AAGGTTAAGATGGATTGAAGAAGAAAGAAACACATAAA- - - - -GGAACAATTAAGCCT
1756 GCAGTATGTTAATTTGATATGAGTCTGTTGGAAACATTTGATACAACTGGATTAAGTAT
1806 TCGATTCTTAGTATTGGAATGTCAGCTGTGATGCTGTGACACAAAGTGAATCTCACT
1816 GCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGCTTACAGCTTGTGTTTGGTCAATCC
1866 TTTCAAGGATTTGAAGAAAGCACTGGAAAGAAAGGTGTTGAGCTTGTGTTGTTCAATCC
1876 TGTAAAGTGAAGTGAAGAAAGCTGAACAAATCGAAGTTTCCAAATCATTTT 1926
1926 TCTTGTGAGGTCATAGAAAGCTTAAAAAAGCAGATGAAGCTTAATGATTT 1976

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 182207
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97750C.1
US-10-425-115-182207

Query Match 26.6%; Score 650.6; DB 20; Length 2651;
Best Local Similarity 60.2%; Pred. No. 6e-164;
Matches 1117; Conservative 0; Mismatches 729; Indels 9; Gaps 2;

QY	121	AGTGCACCAAGTCGAGGTTCCACCGGCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTT	180
DB	517	AGTGCAACAAGGTGGCGCGCGCGCGCGGAGCACCGCGAGCAAGATGAAGGTGAGGGT	576
QY	181	GAGGAGACTTTCTTCCCTCATGACCCCTTTGAGGCAGTTTCAAGAACAAAGCCAGCTTCCAA	240
DB	577	GAAGGAGACCTTTCTTCCCGACGACCGTTCCGGCGTTCAAGGGGAGCGCGCGGGGAC	636
QY	241	GAAGTTCAAGCTTGGGCTTCAGTTCCTTCTTCCCATTTTTCGAATGGCTCCCAAAATACAC	300
DB	637	GCAGTGGCTCATGGCGTCAGGTACTCTTCCCATCTCGGACTGGGTGCGGAGTACTC	696
QY	301	CTTTCAGTTCTTGAAGCTGACCTCATAGTGGCATCACCATCGCTAGCTTGGCCATTCC	360
DB	697	CTTGTGCTCTTCAAGTCGACCTCGTCGGCGCCTCACCAATGGCAGCCTCGCCATTCC	756
QY	361	TCAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTGGACTATATTTCGAG	420
DB	757	TCAGGGCATTAGCTACGCGAAGCTGGCAAGCTTGGCTCCCATTAATCGGGCTGTATTCGAG	816
QY	421	CTTTATACCAACATGATTTATTCGATGATGGTAGTGGTAGCTCAGAGGATTTGGCAGTGGGAC	480
DB	817	CTTTCGTGCGCCGATGGTGTACGCGGTGCTGGGAGCTCCCGTGACCTGGCGGTGGGCC	876
QY	481	TGTGGCGGTGGATCGCTTCTGATGGTTCGATGTTGAGTAAATGCGGTGATCCCAATGA	540
DB	877	GGTGTGATCTCGTCGCTGATCATGGGTCCATGCTGGCGAGCGCGTGGCCGCCACCGC	936
QY	541	AGACCCAAAGCTTTACCTCCACCTGGCTTTCACAGCTACATTAATTTGCTGGTGTATTTCA	600
DB	937	GGAGCCGACGCTGTTCTCGAGCTGGCTTCACTCCACCTGTTTCGCGGGGCTGGTGCA	996
QY	601	GGCTGCTTGGGCTGCTGTTTAGGTTGGGTTGATCGTGGATTTTCTGTCAATGCAACCAT	660
DB	997	GGCCTCCCTGGGCATCTCAGGCTCGCTTCGTCATCGACTTCTGTCTCAAGGCGACGCT	1056
QY	661	AATAGGGTTCACTGGGAGGACGCCAGGTGGTGTCTGCAGCACTAAATCGATTCT	720
DB	1057	GGTGGGGTTTCATGGCGGGCGGCCCATCATCGTGGCGCTGCAGCAGCTCAAGGCGCTGT	1116
QY	721	TGGCCTTGAGCATTTTCAACCCATCGAGCTGATATCATATCAGTATGATGCGCTCTGTTTTAC	780
DB	1117	GGGCATCGTCCATTCCACCAACGAGATGGCATCGTCCCATGTCATGGCCTCCGCTTCCA	1176
QY	781	CCAAACTCATGATGAGGTGGGAAAGTGTGTGTTAGGATGTTCTCAATTTCTTCCT	840
DB	1177	CCACACGAGGAGTGTCTGGGAGACGATCTTCATGGGCGTCTGCTTCTCTTCTTCT	1236
QY	841	CTTTAGCACAGATACTTCAGCAAAAACGACCAAGTTTTTTTGGTGTCTCAGCAATGC	900
DB	1237	GCTGTGGCGAGGCATGTGAGCATCAGATGGCCAAAGCTTTTCTGGGTTTCGGCGTGGCG	1296

QY	901	GCCATTGACGTCGCTTATATTGGGAAGTCTCTTGGTTTATTTTCACTCAACGCGAGAGCA	960
DB	1297	GCCCCTGGCATCGGTCAACCATCTCGACGCTGCTTGTTCCTTCTTCAAGCTCAGAACCA	1356
QY	961	CGGTGTTGAAGTATAGGAGAACTGAAGAAAGGTTTGAATCCACCATCACTCACAATCT	1020
DB	1357	TGGCATCAGCATCATTTGGGCAGCTCAAGTGGCGCTGAATCGCCCTCGTGGGACAAGCT	1416
QY	1021	GGTATTTGTCGCTTACATCACTACAGCTGTCAAACTGGCATTTGTTGGCATCAT	1080
DB	1417	CCTGTTTGACACGCGGTATTTAGGCCTCACCAATGAAGACTGGCCTTGTCAACGGATCAT	1476
QY	1081	ATCACTTGGGAAAGAAATAGCAAGTAAAGCTTTGCAATGTATAAAAAATTACAATAT	1140
DB	1477	CTCACTGACGGAAGGATAGCGGTTGGTAGAACATTTGGCTCACTCAAGGACTACCGAT	1536
QY	1141	TGATGCAACAAAGAGATGATAGCTATTTGGGACCAATGAACGTAAGTGGTCTTTTCACTC	1200
DB	1537	AGATGGAAACAAGGAGATGATGGCCATAGGTTGATGAATGTTGTTGGGTCTCGACATC	1596
QY	1201	TTGCTACCTCAACAAGGACATTTTCGGTTCCGCTGTGAACATTAACGCTGGATCAA	1260
DB	1597	ATGCTACGTAAACAAGGTGCGTTCTCCGCTCTGCTGTAAACCAACACCGCGGTGCAA	1656
QY	1261	GACAGCAGCTTCCAAACATTATATGTCATTCGAGTAATGTTGACATTTGTTTCTCTGAC	1320
DB	1657	GACTGCCATGTCACACGTGATCATGGCGTGACTGTGATGGTCACGCTGCTCTCAT	1716
QY	1321	ACCTTGTTCATATACATCCCTCGTGGTGTATCAGCTATTTATCGTATCTGCAATGCT	1380
DB	1717	GCCACTGTTGTTGTACACACCAACGTTGCTCCGAGCGATCATCATCGCGCGGTGAT	1776
QY	1381	TGACCTCATAGATATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTCT	1440
DB	1777	CGGCTGATGATGATTTCCCGCGGTGTACCACTCTGGAAGATGGACAAGATGATTTCT	1836
QY	1441	GGTGTGATGATGATATCATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAAT	1500
DB	1837	GGTGTGGTTTGGCGGTTTGGCGGCTCATCTTCTCATCTCAGTCCAAGAAGGCTTGGAT	1896
QY	1501	AGCTATTTGATATCTGTACTTTCGGGTACTTCTATTTATTTGCAAGGCCAAGACATTCGT	1560
DB	1897	AGCGGTGGTATATCTATATTTAGGTTGTTGATCGAGATCACGAGGCCCAAGATGATGT	1956
QY	1561	TTTGGGCAACATTCCAAATTTCTGTATATACCGAAATGTTGAGCACTACAAAATGCAAA	1620
DB	1957	TCAAGGGAACATCAAGGGGACTGATATTTTACAGACCTGTCATCAAGAGAGGCCCA	2016
QY	1621	ACATGTTCTGGAATGCTAAATTTAGAGATTGATGACCAATTTTACTTTTGCCTTGGCCAG	1680
DB	2017	AAGATTTCTGGGTTCTTGATCTTGGCCATTGAGCACCGATAAATTCGCCCACTCCAA	2076
QY	1681	CTATTTAAGAGAAAGATCAAGGTGATTTGATGAAGAAAGAAAGAAATTAAGCTAC	1740
DB	2077	CTACCTGAATGAAGAAATTAAGATGGAT-----AGAGGAAGAATCTTTTGAACAGGA	2130
QY	1741	AGGGGAGACTAGTTTGCAGTATGTTATAATTCATATGATGCTGTTGGAAACATTTGATAC	1800
DB	2131	TAAACATCTGAATCCCATTTCAATCTTGGATCTGTGAGCTGTTCTCGCAATTTGACAC	2190
QY	1801	AAAGTGAATAAGTATGCTTTGAAGAGGTGAAGAAATTACAGAGAGAAGAGCTACAGCT	1860
DB	2191	AAAGTGGATAGGTTTCTCTCATTTGACATAAAGAAATCAATAGAGAAACGTTGCTGGAGCT	2250
QY	1861	TGTTTTTGTCAATCTGTGAAGTGAAGTGAAGAACTGAACAAATCGA---AGTTCCA	1917
DB	2251	TGTGTTGTCAATCCAACTGGAGAGTCAATGAGAAAAATACAACTGCAACAGAGCTGA	2310
QY	1918	AAATCATTTAGGGAAGAAATGATCTATCTGACTGTTTGAAGAGGCGGTTTGGAGCA	1972
DB	2311	AAACTATTTTAGGCCAGATTTGCTGTATCTGACCATGCGGAAGCAATCGCTTCA	2365

RESULT 10

US-10-437-963-77005

Sequence 77005, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 77005

LENGTH: 2780

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_76946C.1

US-10-437-963-77005

Query Match 26.4%; Score 647.2; DB 19; Length 2780;

Best Local Similarity 59.8%; Pred.No.5.1e-163;

Matches 1103; Conservative 0; Mismatches 738; Indels 3; Gaps 1;

Qy 125 CACCAAGTCGAGGTTCCACGCCACACGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG 184

Db 457 CACAAGTTCGAGTTCCACCGAAGAAGAGCTCATCGAGGTTACCGAGCGCGTGAAG 516

Qy 185 GAGACTTTCTTCCGTATGACCCCTTTGAGGCGAGTTCAAGAACAAAGCCAGCTTCCAGAAG 244

Db 517 GAGACGTTCTTCGCCGACGACCCGTTGCGGCGAGTACAAGGACGAGCGGATGTCAAAGAAG 576

Qy 245 TTCATGCTTGGCTTCAGTCTTCTTCCCATTTTTCGAATGGGCTCCCAATACACCTTT 304

Db 577 GTGTTGATCAGCTCGAAGACTTCTTCCGGTGCTGAGCTGGGGCGGCGACTACACCTTC 636

Qy 305 CAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGTACGTTGGCCATTCCTCAG 364

Db 637 CGCAAGTTACGGGGCGACCTCGTCTCGGGCTCACCATTGCCAGCTCTGCATACCCAG 696

Qy 365 GGCATCAGTTATGCGAAGCTCGCCAACTCCCTCCAAATTTCTTGGACTATATTCGAGCTTT 424

Db 697 GACATCGGTTATGCGAAGCTTCTGCGCTATTACCAAAATATGGAAGCTGTACAGCAGCTTC 756

Qy 425 ATACCACATTGATTTATGCGATGATGGGTAGCTCGAGGATTTGCGAGTGGGACTGTG 484

Db 757 GTACCGCTTTGATATACGCGATGATGGGAAGTTCCAGGGATATGCCATTTGTGCCAGTG 816

Qy 485 GCGGTTGGATCGCTTCTGATGGTTTCGATGTTGAGTAATGCGGTTGATCCCAATGAAGAC 544

Db 817 GCGGTCGTTTCGCTGCTGCTTGGCACTCTCTCCAGATGAGTTTGTATCCCAAGAAGAT 876

Qy 545 CCAAGCTTTTACCTCCACTGCTTTTACAGCTTACATTAATTTCTGGTGTGTTTTTCAGGCT 604

Db 877 CAGGAAGAGTACACCGGCTGCTTCACTGCAACCTTCTTCTGCTGGGGTCAACCCAGCA 936

Qy 605 GCCTTGGTCTGTTTAGGTTGGGGTTGATCGTGGATTTTCTGTACATGCAACCATATA 664

Db 937 GTGCTCGGATTCCTCAGGCTAGGGTTTATCATAGAGTTCTTGTCCCATGCTGCCATCGTT 996

Qy 665 GGGTTTCATGGGAGGAGCAGCCAGGTTGTGTCGAGCAACTAAATCGATTTCTTTGGC 724

Db 997 GGATTCATGGGGAGCGGCATCACCATTGCCCTTCAGCAGCTTAAAGGCTTCTTTGGA 1056

Qy 725 CTTGAGCATTTTCAACCCATGGAGTGATATCATATCAGTATGATGCGCTCTGTTTTCAACCA 784


```
Db 1683 ATGG---GTGGAAGATGAGGATAACCTGTGCAAGAGCGTCGGGCACGATCTGCAATACTT 1739
Qy 1765 TATAATTGATATGAGTCTGTGTGGAACAATGTATPACAAGTGGAAATAGTATGCTTGAAGA 1824
Db 1740 GGTCTTGTGATCTGTGTGTGTCACCTCTGTGCGAACAATCTGGAGTTCGGATGCTACTAGA 1799
Qy 1825 GGTGAAGAAGATTACAGAGAGAGAGAGCTACAGCTTGTGTTTGGTCAATCCTGTAAGTGA 1884
Db 1800 AGTACACAAGAGCCTCGAAGCAAGAGGGATCACGATAGCTCTGACGAATCCGAGGCTAGA 1859
Qy 1885 AGTGATGAAGAACTGAACAAATC---GAAGTTCCAAAATCATTTAGGGAAGAAATGGAT 1941
Db 1860 GGTAAACAGAGAGCTGTGTCTGTGATAGCTCAGGGACATCTTAGGGATGAGTGGGT 1919
Qy 1942 CTATCTGACTGTTTGAAGAGCGGTGGAGCATGCAACTTCAATCTACGTGCAAGCAAAAC 2001
Db 1920 CTTCTCAGCGTCAAGAGCGCCATCACGGCGTGTGATACGGCGTGCAGATATCCAGAAA 1979
Qy 2002 GAACCCAAAGAAATGAA 2020
Db 1980 TAAGGGAGAAAGCAAGTA 1998
```

RESULT 12

```
US-10-437-963-73410/c
; Sequence 73410, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)/B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 73410
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73694C.1
US-10-437-963-73410
```

```
Query Match 23.3%; Score 571.4; DB 19; Length 2429;
Best Local Similarity 58.1%; Pred. No. 1.3e-142;
Matches 1075; Conservative 0; Mismatches 726; Indels 48; Gaps 2;

Qy 123 TGCACCAAGTCGAGTTCACCGCCACAGCGGTTTTCAAGTCTCTAAAGTACTCTTTGA 182
Db 2260 TGCACAAAGTGTGCTGCCGAGAGGAGGTGCGACGGCGAAGGGCGCTGAGGCAGCGGTGG 2201
Qy 183 AGGAGACTTTCTCCCTGATGACCCCTTTGAGGCGATTCAGACACAGCCAGCTTCCAGA 242
Db 2200 CGGAGGTGTTCTTCCCGCAGACCCGCTCCACCAAGTTTCAAGAACCAAGTCTCGCGCGGC 2141
Qy 243 AGTTCATGCTTGGCTTTCAAGTCTTCTTCCCAATTTTCGAATGGGCTCCCAAAATACACCT 302
Db 2140 GGTGTTGCTCGCGCTGAGTACTTCTTCCCATCTTCCATTTGGGGTCCGACTACAGCC 2081
Qy 303 TTCAAGTTCTTAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTC 362
Db 2080 TCCGCTCTCCGCTCCGAC----- 2061
Qy 363 AGGCATCAGTTATGCCAAGCTGCGCAACCTCCCTCCAATTTTGGACTATATTCGAGCT 422
```

```
Db 2060 --GGAATCAGCTAGCGCAAGCTCGCCAACTGCCTCCAATCATTTGAGCTATATTCGAGCT 2003
Qy 423 TTATACCACCACTTGAATTTATTCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTG 482
Db 2002 TCGTGGCGCGCTGATCTACTCTTGTGGGTAGCTCGCGGATCTGGCTGTAGGGCGCG 1943
Qy 483 TGGCGGTTGGATCGCTTCTGATGGGTTCGATGTTGATGTAATCCGTTGATCCCAATGAAG 542
Db 1942 TGTGCGATAGCGTCCGTGGGTGATGGGGTCAATGCTCCGGCAGGCGGTGTCCCGGACAGG 1883
Qy 543 ACCCAAAGCTTTACCTCCACCTGGCTTTACAGCTACATTAATTTGCTGGTGTGTTTTTCAGG 602
Db 1882 AGCCCATCTCTACCTCCAGCTCGCTTCACTCCACCTTCTTCGCGGGGTCTTCCAAG 1823
Qy 603 CTGCTTGGGTCTGTTTAGTGGGTGATCGTGAATTTTGTGCACATGCAACCAATAA 662
Db 1822 CTTCCCTCGCTTCTCTCGGCTGGGGTTCATCGTGAATTTCTTGTGGAAGCGACGCTGA 1763
Qy 663 TAGGGTTTCATGGAGAGCAGCCACGTTGGTGTCTGCAGCAACTAAATTCGATTTCTTG 722
Db 1762 CGGGATTCATGGCGCGCGCGGATCATCGTGTGCTGCAGCAGCTCAAGGGATTTGCTG 1703
Qy 723 GCCTTGAGCATTTACCCATGGAGCTGATATCATATCAGTGTGCGTCTCTGTTTTCACCC 782
Db 1702 GGATCATCTTTCACGTGCGAGATGGGGTTCGTCCAGGTGATGCACTCGTCTTCAAGC 1643
Qy 783 AAATCATGAGTGGAGGTGGGAAAGTGTGTGTTAGGATGTGTTTCATTTTCTTCTCTC 842
Db 1642 ACCACGAGTGGGGGTGGCAGACCATCTCATGGCGGTGCGCTTCTCGCGCTCTCTC 1583
Qy 843 TTAGCAAGAATACCTTCAGCAAAAACAGCAAGGTTTTTTGGGTGTGAGCAATGGCGC 902
Db 1582 TCACCAACAGCCACATCAGCGCCAGGAACCCAAAGCTTTTCTGGGTGTCTAGCAGCTGCTC 1523
Qy 903 CATTGACGCTCCGTTATATTGGGAAGTCTCTTGTGTTTATTTCACCTCAGCCGAGAACGACG 962
Db 1522 CACTGACATCAGTGTATCTCTACCATCATCTCGTTCGTCCAGCAAGC-----TCATG 1469
Qy 963 GTGTTGAAGTGTAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATACTGG 1022
Db 1468 GCATCAGTGTATTGGCGATCTCCCAAGGGATTGAACCCCTCTCTCAGCGAATATGCTGA 1409
Qy 1023 TATTGTGCGCTTACATGACTACAGCTGTCAAACTGGCATTTGTTGGCATCATAT 1082
Db 1408 CTTTCAGTGGCTCTCTAGTAGGACTCGGCTTAAACACAGGAATTAAGACTGGCATACTAT 1349
Qy 1083 CACTTCGGGAAGGAATAGCAGTAGGAAGAGCTTTTGCATGTATAAAATTTACAATATTG 1142
Db 1348 CTCTCACTGAAGGATAGCAGTAGGAGGACATTTGCATCCATCAACACTACAGGTTG 1289
Qy 1143 ATGGCAACAAAGATGATAGCTATTGGGACCAATGAACGCTAGTTGGTCTTTTCACTCTT 1202
Db 1288 ATGGAAACAAGGAGATGATGGCCATTTGGTGTCTATGAACATGGCTGGCTCATGTGCTCCT 1229
Qy 1203 GCTACCTCACAACAGACCACTTTTCGGGTTCGGCTGTGAACATAAGCTGGATGCAAGA 1262
Db 1228 GCTATGTCCACAGGATCGTTCTCGAGGTTCGGCGGTGAACCTACAGCGGGGTGCAAGA 1169
Qy 1263 CAGCAGCTTCCAACTAATAATGTCTACTTGCAGTAATGTTGACATTTGTTATTCTCTGACAC 1322
Db 1168 CGCGGTGTGAACATCTGATGCGCGTGTGCTGAGCGTGTGCTGCTGCTGCTGCTGATGC 1109
Qy 1323 CTTTGTTCATATACACTCCCTCGTGGTGTGCTATCAGCTATTAATCGTATCTGCAATGCTTG 1382
Db 1108 CGCTGTTCCACTACACCCCAACGTCATCTCTCGCGGATCATCATCACCGCGGTGATCG 1049
Qy 1383 GACTCATAGATTATGAAGCAGCCATCCATCTATTAAAGTTGACAAATTTGACTTTGTGG 1442
Db 1048 GCCTCATCAGCGTCCGGGGCGCGCAGGCTGTGGAAGGTGCGAAGCTCGACTTCTCTCG 989
Qy 1443 TGTGCAATGAGTCATACATTTGGCGTGTCTTTGGCAGTGTGTAATTTGGCTTGTAGTCATAG 1502
Db 988 CTTGCATGGCGCGCTTCTCTCGGCGTCTCTCTGCTCCGCTCCAGATGGGCTCGCCATCG 929
```

```
QY 1503 CTATTGTAATATCTGTACTTTCGGGTACTTCTATTATTGCAAGGCCCAAGGACATTCGTTT 1562
Db 928 CCGTCGGCACTCCCTCTTCAAGATCTCTCCAGGTCAACCCGCCCAAAATGGTGTCA 869
QY 1563 TGGGCAACATTCAAATTCCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAAC 1622
Db 868 AGGCGCTCGTCCCGGCAACCGGAGCTACCGGAGCATGGGCGAGTACAGGGAGGCCATGC 809
QY 1623 ATGTTCTCGAATGCTAAATCTTAGAGATTGATGCACCAATTTACTTTTGCCAATGCCAGCT 1682
Db 808 GTGTGCCGTCTTCTCTCGTCTCGGCGTCAGTCCGCACTACTTCTGCGCAACTCCATGT 749
QY 1683 ATTAAAGAGAAGGATCACCAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACAG 1742
Db 748 ACCTCGCGAGAGATCATGAGGTTCTCGCGGAGNAGACGAGCGGCCCAAGTGCA 689
QY 1743 GGGAGACTAGTTTGCAGTAGTTATTAATGATATGAGTGTGTTGAAACATTTGATACAA 1802
Db 688 ACCAGTGCCTGTGAGATGCATCATCTCGACATGAGTGTGTTGAGCGATCGACACGA 629
QY 1803 GTGMAATAAGTATGCTTGAAGAGGTGAAGAGATTACAGAGAAAGAGACTACAGCTTG 1862
Db 628 GTGCGCTCGATGCACTTTCAGAGCTGAAGAAAGTGTGCGAAAGAAAGAAACATCGAGCTTG 569
QY 1863 TTTTGTCTCAATCCTGTAAGTGAAGTGATGAAGAAACTGAAACAAATCGAAGTTTCCAAATC 1922
Db 568 TGCTGGCCAAACCGGTTGGATCGGTGACCGAGAGGCTGTACAACTCGGTGTCGGCAGA 509
QY 1923 ATTTAGGAAAGAAATGATCTATCTGACTGTTGAAGAGGCCGTTGGAGC 1971
Db 508 CGTTCGGCTCAGACCGCGTGTCTTTCAGCGTCGCCGAGCGCTCGCGGC 460

RESULT 13
US-10-425-114-15719
; Sequence 15719, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15719
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-059-F10_FLI
US-10-425-114-15719

Query Match 23.2%; Score 567; DB 18; Length 2656;
Best Local Similarity 57.0%; Pred. No. 2.1e-141;
Matches 1060; Conservative 0; Mismatches 795; Indels 6; Gaps 1;

QY 123 TGCAACCAAGTCGAGGTTCCACGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGA 182
Db 300 TGCAACCAAGTGTCTGTCGCCGAGCGGGGACGACGGGAAAGGCGCTGCGGACGGCCTGG 359
QY 183 AGGAGACTTTCTTCCCTGATGACCTTTTGAGGAGTTCAAGAACAAAGCCAGCTTCCAAGA 242
Db 360 CCAGGAGTGTCTTCCCGGACGACCCGCTGCACCAAGTTCAAGAACCAAGTGTGTCGGCGCGC 419
QY 243 AGTTCAATGTTGGCTTCAGTCTTCTTCCCAATTTTCGATGGGCTCCCAATACACTT 302
```

```
Db 420 GCCTCGTGTGCGCTGCACTACTTCTTCCCACTTCTCCAGTGGGGTCCGCCCTACAGCC 479
QY 303 TTCAGTTCTTGAAGCTGACCTCATAGCTGGGATCACCATCGCTAGCTTGGCCATTCTCTC 362
Db 480 CGGCGCTCTCTGGCTCCGACCTCGTCCGCGCTCACCATTGCCAGCTCGCCATCCCGC 539
QY 363 AGGCAATCAGTTATGCAAGCTCGCAACCTCCCTCCAAATCTTCTTGACTATATTTGAGCT 422
Db 540 AGGCAATCAGCTACGCCAAGCTCGCAACCTCCCGCAATCGTTGGCTTATATTTCCAGCT 599
QY 423 TTATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 482
Db 600 TGTGCGCGCTCTCATCTACGCGCTGCTGGGAGCTCGCGGACCTCGCGGCTGGGGCCGG 659
QY 483 TGGCGGTTGATGCTGCTTCTGATGGGTTGATGTTGATGATGATGATGATGATGATGATG 542
Db 660 TGTCCATCGCTCGCTGCTGATGGGTTGATGATGATGATGATGATGATGATGATGATGATG 719
QY 543 ACCCAAGCTTTTACCTTCCACCTGGCTTTTCCAGCTTACATTTTGTGCTGTTTTCAGG 602
Db 720 AGCGCTCTCTTACCTGACCTGGCTTTCACCGCACCTTCTTCGCGCGCTCTTCCAGG 779
QY 603 CTGCTTGGCTGCTGTTTGTGTTGGGTTGATCGTGGATTTTCTGTACATGCAACCATAA 662
Db 780 CGTCCCTGGGATTTCTTCAGGCTGGGCTTCATCGTGGACTTCTGTCCAAAGGCGACGCTGA 839
QY 663 TAGGTTTCATGGGAGGAGCAGCAGCTGCTGTGCTGAGCAACTTAAATCGATTCTTG 722
Db 840 CGGCTTTCATGGCGCGCGCGCTCATCGTGTGCTGAGAGCTCAAGGCGCTGCTCG 899
QY 723 GCCTTGAGCATTTTCAACCATGAGCTGATATCATATCAGTGATGCGCTCTGTTTTCACCC 782
Db 900 GCATCTCCCACTTCACTCCCATGGGATTCTCGACGTCATGGCTCGTCAAC 959
QY 783 AAACCTCAGTGGAGTGGGAAAGTGTGTTAGGATGTGTCTTCACTTTTCTTCTCTCC 842
Db 960 GCCACGACGAGTGGAAAGTGGCAGACGATCGTATGGGCTCCGCTTCTTCGCACTCTCC 1019
QY 843 TTAGCACAAAGTACTTTCAGCAAAACACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGC 902
Db 1020 TCCTCAGCGCCAAATCAGCGCCAGGACCCAAAGCTTTTCTGGGTATCAGCAGGTGCTC 1079
QY 903 CATTCAGCTCCGTTATTTGGGAAGTCTCTTGTTTATTTTCACTCAGCGCGAGAAAGCAG 962
Db 1080 CCTCGCGTGGTGTATCTCCACCTCTCTCTTCTTCTTCTGAAATCCCCAGCATCA 1139
QY 963 GTGTTGAAGTATAGGAGAACTGAAGAGGTTTGAATCCACATCAGTCAACAATCTGG 1022
Db 1140 GTGTT-----ATTGGCATCTCTCCCGAGGGAGTGAACCTCTTTCGGCGAATGCTCA 1193
QY 1023 TATTTGTGTCGCTTACATGACTACAGCTGTCAAACTCGCAATGTCGTTGGCATCATAT 1082
Db 1194 GCTTACGGGCTCTTATGTGGCGCTGACATCAAAACCGGGATCATGACAGGATCCTGT 1253
QY 1083 CACTTCGGAAGGAATAGCAGTAGGAAGAGCTTTTGCAATGTATAAAATTTACAAATTTG 1142
Db 1254 CTTTAAACGAAGAGGATCGCAGTGGGAGGACCTTCGCGTCCATCAACAACCTACAGGTG 1313
QY 1143 ATGGCAACAAAGATGATAGCTATTTGGACCATGAACGTAGTTGGTTCTTTCACCTCTT 1202
Db 1314 ACGGGAACAAAGAGATGATGAGCGATCGGCTGATGAACATGCGGGGCTCTCGCGCTCT 1373
QY 1203 GTTACTCTCAACAGAGCACTTTTTCGCTTGGCTGTGAACCTATAACGCTGGATGCAAGA 1262
Db 1374 GTTACGTGACAGAGGGTCTTCTCCCGTTCGCGGTGAACTACAGCGCGGCTGACAGA 1433
QY 1263 CAGCAGCTTCCAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1322
Db 1434 CGGCGCTGTCCAAACGTGTGATGGCGCGCGGTGCTGGTACGCTGTGTTCTCTCATGC 1493
QY 1323 CTTTGTTCATTTACACTCCCTGCTGCTATCAGTATTTATCGTATCTGCAATGCTTG 1382
Db 1494 CGCTGTTCCACTACACCCCGAAACGTGATCTCTGGCGCGGATCATCATCACGGCGGTGG 1553
```

Qy	1393	GACTCATAGATTATGAAGCAGCGCATCCATCTATTAAAGTTGACAAATTTGACTTTGTGG	1442
Db	1554	GGCTGGTGGACGTGCGCGCGCGCGAGCTGTGGAAGGTGGACAAGCTTGGACTTCTCTGG	1613
Qy	1443	TGTGCATGAGTGCATACATTTGGCGTGGTCTTTTGGCAGTGTTCGAAATTTGGCTTAGTCATAG	1502
Db	1614	CGTGGCTGGCGGCGTTCCTCGCGTGTCTCGTGTCCGTGCAGACGGGCCCTGGGGGTGG	1673
Qy	1503	CTATTGTAATCTCTACTTCGGGTACTTCTATTATTATTCGAAGGCCAAGGACATTCGTGTTT	1562
Db	1674	CCGTGGGATCTCGCTCTTCAAGTTCCTGCTCGAGTCAACCGGCCCAACGTCGTGGTGG	1733
Qy	1563	TGGGCAACATTCGAAATTCGTGATATACCGAAATGTTTGAGCACATCAAAATGCAAAAC	1622
Db	1734	AGGGCTCGTCCGGGAGCGAGACTACCGCAGCGTGGCGCAGTAGTACCGGAGGCGCTCC	1793
Qy	1623	ATGTTCTTGGAAATGCTAATTTTAGAGATTGATGCACCAATTTACTTTTGCAATGCCAGCT	1682
Db	1794	CGGTGCCGGGCTTCCTCGTCTGGCGGTGAGTCCGCGCTCTACTTTGCGCAACTCCATGT	1853
Qy	1683	ATTTAAGAGAAAGGATCACAGGTGGATTGATGAAGAAGCAAGAAATTAAGCTACAG	1742
Db	1854	ACCTGGTGGAGGGGTCAATCGCTACCTCCGCGCAGAGGAGAGCGCGCTCAAGTCGA	1913
Qy	1743	GGGAGACTGATTTGCAGTATGTTATTAATGATATGATGCTGCTGTGGAAACATTTGATACAA	1802
Db	1914	ACCACCCCTCCATCCGATGCGTCTGCTCGACATGGGCGCGCTCGCGGCGATCGACACGA	1973
Qy	1803	GTGGAATAAGTATGCTTGAAGGTGAAGAGATTTACAGAGAGAGAGAGCTACAGCTTG	1862
Db	1974	CGGTCTAGACGCTGTCCGAGCTCAAGAAAGTCTTGCAAAAAGAAACATCGAGCTGG	2033
Qy	1863	TTTTGGTCAATCCTGTAAAGTGAAGTGAAGAAAACGAAACAAATCGAAGTTTCCAAAATC	1922
Db	2034	TGCTTGCCAAACCGGTGGGGTCCGTGGGCGAGAGGATGTTCAACTCGGCGGTGGGCGAGA	2093
Qy	1923	ATTTAGGGAAGAAATGGAATCTATCTGACTGTTTGAAGAGCGCGTTCGAGCATCCAACTTCA	1982
Db	2094	GCTTCGGTTCGGGCGCCTCTTCTTCAGGTAGCGGAGGCGCGTTCGCGGCGGGCGTGGCA	2153
Qy	1983	A	1983
Db	2154	A	2154

RESULT 14

```

US-10-425-115-75763
; Sequence 75763, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 75763
; LENGTH: 2713
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_169123C.1
US-10-425-115-75763

```

Query Match	23.2%;	Score 567;	DB 20;	Length 2713;
Best Local Similarity	57.0%;	Pred. No. 2.1e-141;		
Matches 1060;	Conservative	0;	Mismatches 795;	Indels 6

Best local similarity 37.0%; Fied. NO. 2.1E-141;
Matches 1060; Conservative 0; Mismatches 795;
Indels 6; Gaps 1;

Qy	123	TGCAACAAAGTCGAGTTTCCACCGCCACAGCGGTTTTTTTCAAGTCTCTAAAGTACTCTTTTGA	187
Db	357	TGCAACAAAGTGTCTGCTCCGAGCGCGGAGCGA	416
Qy	183	AGGAGACTTTCTTCCCTGATGACCCTTTGGGCGAGTTTCAAGAACAAAGCCAGCTTCCAAGA	242
Db	417	CCGAGGTGTCTTCCCGAGCGACCCGCTGCACAGTTCAGAACCAAGTCGTGCGGCGCGC	476
Qy	243	AGTTTATGCTTGGCCTTCAGTTCTTCTTCCCATTTTTTCGAATGGGCTCCCAAAATACACCT	302
Db	477	GCCTCGTGGCGCTGCATCTACTTCTTCCCATCTTCCAGTGGGGTCCGCTACAGCC	536
Qy	303	TTCAAGTTCTTGAAGCTGACCTCATAGCTGGCATCAACATCGCTAGCTTGGCCATTCCTC	362
Db	537	CGCGCTCTCTCGGCTCCGACCTCGTCCGCGCCTCACATTTGCCAGACTCGCCATCCGCG	596
Qy	363	AGGCGATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTTGGAGTATATTTTCGAGCT	422
Db	597	AGGGAATCAGCTACGCCAAGCTCGCCAACTTCCGCGCAATCGTTTGGCTATATTTCCAGCT	656
Qy	423	TTATACCAACCATTAATTTATTCGATGATGGGTAGCTCGAGGGATTTTGGCAGTGGGGACTG	482
Db	657	TCGTGCGCGCTCATCTACGCGCTGTGGGAGCTCGCGGACCTCGCGGTGGGCGCGG	716
Qy	483	TGGCGGTTGGATCGCTTCTGATGGGTTGATGTTAGTAAATGCGGTTGATCCCAATGAAG	542
Db	717	TGTCATCGCGTCTGCTGGTATGGGTCATGCTCCGGACGCGGTCGTCCGCGACGAGC	776
Qy	543	ACCCAAAGCTTTACCTTCCACTGCTTTCACAGCTACATTAATTTGCTGGTGTTTTTTCAGG	602
Db	777	AGCGCTCTCTTACCTCGACCTGGCCTTTCACCGCACCTTCTTCGCGCGGCGGTCTTTCAGG	836
Qy	603	CTGCTTTGGGCTCTGTTTAGGTGGGGTTGATCGTGGATTTTCTGTACATGCAACCATAA	662
Db	837	CGTCCCTGGGATTTCTTCAGCTGGGCTTCATCGTGGACTTCTGTCCAAGCGACGCTGA	896
Qy	663	TAGGGTTATCGGAGGAGCAGCCACGGTGTGTGTCTGACGCAACTAAAATTCGATTTCTTG	722
Db	897	CGGGCTTCATGGGCGGCGCGCGCTCATCTGTGTGCTGACAGAGCTCAAGGGCGCTGCTCG	956
Qy	723	GCCTTGAGCATTTTCAACCATGGAGCTGATATCATATCAGTGTATGCGGCTCTGTTTTCAACC	782
Db	957	GCATCTCCCACTTCACTCCCAATGGGATTCCTTCGACGTGATGCGCTCGTGTGTCGAACC	1016
Qy	783	AAACTCATGTGAGGTGGGAAAGTCTGTGTTAGGATGTCTTCATTTTCTTCTCTCC	842
Db	1017	GCCACGACGAGTGGAGTGCAGACGATCGTCATGGGCTCGGCTTCTCGCCATCTCTC	1076
Qy	843	TTAGCAAGAATACTTCAGCAAAAACGACCAAGTTTTTTTGGGTGTACGCAATGGCGC	902
Db	1077	TCCTCAGCGCCAAAATCAGCGCCAGGAAACCAAGCTTTTCTGGTATCAGCAGGTGCTC	1136
Qy	903	CATTGACCTCGGTTATATTGGGAGTCTCTTGGTTTATTTTCACTCAGCGCGGAGNAGCAG	962
Db	1137	CCCTGGGCTCGGTGATCATCTCCACCATCTCTCTTTCATCTGGAATCCCGCCAGCATCA	1196
Qy	963	GTGTTGAAGTATAGGAGAACTGAAGAAGGGTTTGAATCCACCATCACTCACAATCTCGG	1022
Db	1197	GTGTT-----ATTGGCATCTCCCGAGGGAGTGAACCTCTCTTGGCGGAACATGTCTCA	1250
Qy	1023	TATTTGTGCGCCTTACATGACTACAGCTGTCAAAACTGGCATTTGTCTTGGCATCATAT	1082
Db	1251	GCTTCAGCGGCTCCTATGTGCGCTGACGATCAAAACCGGATCATGACAGGCATCTCTGT	1310
Qy	1083	CATTGGGGAAGGAATAGCAGTATAGGAAGAAGCTTTTGCATGTATAAAAATTACAAATTG	1142
Db	1311	CTTTAAACAGAAGGGATCGCAGTGGGAGGACCTTTCGCGTCCATCAACAACCTACCAAGTGG	1370
Qy	1143	ATGCAACAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCCTTTCACCTCTT	1202
Db	1371	ACGGGAACAAGGAGATGATGGCATCGGGCTGATGAACATGGCGGGCTCTCTCGCCTCTCT	1430
Qy	1203	GCTACCTTCAACAAGGACCAATTTTCGGTTTCGGCTGTGAACACTATAACCGCTGGATGCAAGA	1262

Db 1431 ||||| GCTACGTCAGACGGGGCTCTCTCCGGTCGGCGTGAACATACAGCGGGCTGCAGGA 1490
Qy 1263 CACGAGCTCCACATTAATGCTACTTCAGTAATGTTGACATGTTATTCCTGACAC 1322
Db 1491 CGGCGTGTCCAAACGTCGTGATGGCGCGCGTCTGGTGACGCTGCTTCTCATGC 1550
Qy 1323 CCTTGTTCATTAACATCCCTCGTGTCTATCAGCTATTAATCGTATCTGCAATGCTTG 1382
Db 1551 CGCTGTTCCACTACACCCGAACTGATCCTGGCGGGATCATCATCAGCGCGTGGTG 1610
Qy 1383 GATCTAGATTAAGACAGCCATCCATCTATTAAAGTTGACAAATTTGACTTTGTGG 1442
Db 1611 GGTGTGTGAGCTGCGCGCGCCGCCAGGCTGTGGAAGTGGACAAAGCTGGAATTCCTCG 1670
Qy 1443 TGTGCATGAGTGACATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGCTAGTATAG 1502
Db 1671 CGTGCCTGGCGGCTTCTCGGCGTGTCTGTGTGTCGGTGACAGCGGCGTGGCGTGC 1730
Qy 1503 CTATTGTAATATCTGTACTTCGGGTACTTCTATTATTGCAAGGCCAAGACATTCGTTT 1562
Db 1731 CCGTCGGCATCTCGCTCTCAAGTCTCTGTCAGGTCAACCGCCCAACGTCGTGGTG 1790
Qy 1563 TGGGCAACATTCGATGATATACGAAATGTTGAGCACTATCAAAATPGCAAAAC 1622
Db 1791 AGGCGCTCGTCCCGGGGACGACAGCTACCGCAGCTGGCGCAGTACCGCGAGCGCTCC 1850
Qy 1623 ATGTTCTCGAATGCTAATCTTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAGCT 1682
Db 1851 GCGTGGCGGCTTCTCGTGTGGCGTGTGAGTCCGCGCTACTTTCGCCAATCCATGT 1910
Qy 1683 ATTTAAGAGAAGGATCACAAAGTGGATTGATGAAGAAGAAAGAAATTAAGCTACAG 1742
Db 1911 ACCTGGTGGAGCGGTCATCGCTACTCCGCGACGAGGAGGAGCGCGCTCAAGTCCA 1970
Qy 1743 GGAGACTAGTTTGCAGTATGTTAATTAATGATGATGCTGTGTGAAACATTTGATACAA 1802
Db 1971 ACCACCCCTCCATCCGATGGTCTCTCGACATGGCGCGCTGCGGGCGATCGACACGA 2030
Qy 1803 GTGGAATAGTATGCTTGAAGAGTGAAGAGTACAGAGAGAGAGAGTACAGCTTG 1862
Db 2031 CGCGTCTAGACGCGTGTCCGAGCTCAAGAAAGTCTCGACACAAAGAAACATCGAGCTGG 2090
Qy 1863 TTTTGTCTAATCTGTAAGTGAAGTGAAGAAACTGAACAAATCGAAAGTTCGAAATC 1922
Db 2091 TGCTTGCCAAACCGGTGGGTGGTGGCGGAGAGGATGTTCACTCGGCGGTGGCGGAGA 2150
Qy 1923 ATTTAGGAGAAATGATCTACTGCTGTTGAAGAGCGGCTTGGAGCATGCAACTTCA 1982
Db 2151 GCTTCGGGTGGGCGCGCTTCTTTCAGCGTAGCGGAGCGCGTGGCGGGCGGTGCA 2210
Qy 1983 A 1983
Db 2211 A 2211

RESULT 15
US-10-424-599-32308
; Sequence 32308, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 32308
; LENGTH: 536

; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129179C.1
US-10-424-599-32308
Query Match 21.2%; Score 520.2; DB 18; Length 536;
Best Local Similarity 98.5%; Pred. No. 3.3e-129;
Matches 525; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 7 GCTAGCTCGCAGTAAAGTTATATAACACATATTTGCTTGTAGAAATACTATTATTCGA 66
Db 4 GCTAGCTCGCAGTAAAGTTATATAACACATATTTGCTTGTAGAAATACTATTATTCGA 63
Qy 67 AGATATGGGAGTGTAGATTATAGTACCTTTGGGCATGAACAACTTTGAGAGAGTGA 126
Db 64 AGATATGGGAGTGTAGATTATAGTACCTTTGGGCATGAACAACTTTGAGAGAGTGA 123
Qy 127 CCAAGTTCGAGGTTCCACCGCCACAGCCGTTTTCAAAGTCTCTAAAAGTACTCTTTGAAGGA 186
Db 124 CCAAGTTCGAGGTTCCACCGCCACAGCCGTTTTCAAAGTCTCTAAAAGTACTCTTTGAAGGA 183
Qy 187 GACTTTCTTCCCTGATGACCTTTGAGGAGTTCAGAACAGCCAGCTTCCAAAGAGTT 246
Db 184 GACTTTCTTCCCTGATGACCTTTGAGGAGTTCAGAACAGCCAGCTTCCAAAGAGTT 243
Qy 247 CATGCTTGGCCCTTCAAGTCTTCTTCCCAATTTTGAATGGGCTCCCAATAACACCTTTCA 306
Db 244 CATGCTTGGCCCTTCAAGTCTTCTTCCCAATTTTGAATGGGCTCCCAATAACACCTTTCA 303
Qy 307 GTTCTTGAAGCTGACCTCATAGCTGGCATCAACCATCGCTAGCTTGGCCATTTCTCAGGG 366
Db 304 GTTCTTGAAGCTGACCTCATAGCTGGCATCAACCATCGCTAGCTTGGCCATTTCTCAGGG 363
Qy 367 CATGAGTTATGCAAGCTGCGCAACCTCCCTCCAAATTTCTTGGACTATATTTCAGGCTTTAT 426
Db 364 CATGAGTTATGCAAGCTGCGCAACCTCCCTCCAAATTTCTTGGACTATATTTCAGGCTTTAT 423
Qy 427 ACCACCATTTGATTTATGCGATGATGGTAGCTCGAGGGATTTGGCAGTGGGACCTGTGGC 486
Db 424 ACCACCATTTGATTTATGCGATGATGGTAGCTCGAGGGATTTGGCAGTGGGACCTGTGGC 483
Qy 487 GGTTCGATCGCTTCTGATGGGTTTCGATGTTAGTAAATGCCGTTGATGCCAATG 539
Db 484 GGTTCGATCGCTTCTGATGGGTTTCGATGTTAGTAAATGCCGTTGATGCCAATG 536
Search completed: August 31, 2005, 09:39:23
Job time : 2775 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 02:28:41 ; Search time 7917 Seconds
(without alignments)
11774.586 Million cell updates/sec

Title: US-10-762-049-17
Perfect score: 2449
Sequence: 1 gcacggctagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	938.4	38.3	2158	3	CNS0A7F1	BX822461 Arabidops
2	670.4	27.4	1959	9	CL961102	CL961102 OsIFCC005
3	655.6	26.8	854	6	CB892638	CB892638 EST645430
4	653.8	26.7	2279	3	AV105934	AV105934 Zea mays
5	605.4	24.7	2001	9	CL959635	CL959635 OsIFCC003
6	581.4	23.7	583	2	AW598447	AW598447 sj91904.y
7	570.8	23.3	1812	9	CL969864	CL969864 OsIFCC019
8	567	23.2	567	4	BM093232	BM093232 saj06610.
9	520.4	21.2	522	4	BG406552	BG406552 sac33f11.
10	516.4	21.1	594	7	CO982632	CO982632 GM89016A2
11	512.6	20.9	2221	3	CNS0A1LZ	CO982632 Arabidops
12	501.4	20.5	828	7	CV130479	CV130479 X39P02a03
13	493.2	20.1	498	4	B1972185	B1972185 sag88c07.
14	476	19.4	982	7	CK269206	CK269206 EST715284
15	468.2	19.1	583	5	BU090662	BU090662 su07b07.y
16	461.4	18.8	489	2	BE804764	BE804764 ss45d02.y
17	459.6	18.8	2321	3	CNS0A1LZ	BE804764 Arabidops
18	448.4	18.3	831	7	CO107436	CO107436 GR_EB003
19	448	18.3	459	5	BU544910	BU544910 GM880003B
20	437.2	17.9	768	7	CN190878	CN190878 UCRCS06.0
21	433.8	17.7	816	6	CB624151	CB624151 OSTIEa12G
22	423	17.3	706	7	CO102483	CO102483 GR_EB002
23	416.6	17.0	738	5	BU025764	BU025764 QHGL1F06.
24	411.8	16.8	728	5	BU025129	BU025129 QH7N01.Y

RESULT 1

CNS0A7F1

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Reference

Authors

Journal

Comment

Features

Source

Location/Qualifiers

1. .2158

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

25	407.2	16.6	652	6	CB035946
26	400.6	16.4	783	7	CK256297
27	398.6	16.3	761	5	BQ851246
28	395	16.1	745	5	BQ996962
29	391.8	16.0	555	2	AW686983
30	390.2	15.9	749	1	AJ794089
31	388	15.8	812	5	BQ796959
32	377	15.4	699	5	BQ849576
33	376.6	15.4	683	5	BQ970363
34	376	15.4	697	5	BQ850117
35	368	15.0	648	6	CB921892
36	365.8	14.9	728	2	AW036000
37	359.4	14.7	844	7	CK269207
38	358.6	14.6	854	5	BQ798101
39	357.8	14.6	673	7	CV474471
40	357.8	14.6	775	7	CK256298
41	357	14.6	847	5	BQ796985
42	356.6	14.6	1534	5	BU693126
43	353.8	14.4	735	7	CN189358
44	352.2	14.4	629	5	BU024884
45	350.4	14.3	675	7	CK901981

ALIGNMENTS

CNS0A7F1 2158 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTFB32C02 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
BX822461 GI:42465936
HTC; GSUT cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 2158)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 2158)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1. .2158
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

		/clone="GSLTFB32C02" /tissue_type="Flowers and buds" /plasmid="pCMVSPORT 6" complement(1..2158) /gene="At3g51895"			
		Query Match 38.3%; Score 938.4; DB 3; Length 2158; Best Local Similarity 70.3%; Pred. No. 4.8e-234; Matches 1328; Conservative 0; Mismatches 551; Indels 9; Gaps 5;			
gene		ORIGIN			
QY	125	CACCAAGTCAGGTTCCACCGCACAGCCGTTTTCAAGTCTCTAAAGTACTCTTTGAAG	184	1016	TCCTTTCTTTACATT-CATGTCCACAGCTGTGCACACCTGGCCACATCACTGGCATCATTTG
Db	119	CACACGGTGGAGGCTCCACAACCTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAG	178	1083	CACCTTGGCGAAGGAATAGCAGTAGGAAGAGCTTTGGCAATGTATAAAAATTTACAATATTG
QY	185	GAGACTTCTCTCCCTGATGACCCCTTTGAGGCACTTCAAGAACAGCCAGCTTCCAGAGAG	244	1075	CTCTCGCTGAAGAGTAGCAGTGGGAGGAGTTTTCGATGTTCAAGAACTACAACATAG
Db	179	GAAACTCTGTTTCCAGACGACCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAA	238	1143	ATGGCAACAAAGAGATGATAGCTATTGGGACCAATGAACGTAGTTGGTTCTTTCCACTCTT
QY	245	TTCAATGCTTGGCCTTCAGTTCCTTCTTCCCAATTTTGAATGGGCTCCCAAAATACACCTTT	304	1135	ACGGGAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCTTTCCATCTT
Db	239	TTTGTGTAGGCTTCAATACCTTCTCCCGATTTTCGAATGGGCAACCGCTACAATCTC	298	1203	GCTACTCAACAGGACCACTTTTCGCGTTTCGCGTGTGAACTATACGCTGGATGCAAGA
QY	305	CAGTCTTGAAGCTGACCTCATAGTGGCATCACCATGCTAGCTTGGCCATTTCTCTAG	364	1194	ACGGGAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCTTTCCATCTT
Db	299	AGTTCCTTCAAAATCAGATCTCATCGCGGAATCACCATCGCTAGTCTCGCCATCCCTCAG	358	1262	GCTACTCAACAGGACCACTTTTCGCGTTTCGCGTGTGAACTATACGCTGGATGCAAGA
QY	365	GGCATCAGTTATGCCAGTCGCCAACCTCCCTCCAAATCTTGGACTATATTCGACTTT	424	1254	GTTCACCTCAACCGGACCACTTTCAAGGTTCGGCAGTGAACACACGCGGGTTCGAAGA
Db	359	GGCATCAGTTATGCCAACCTTGTAACTTGGCCCCAAATCTTGGCCTTTATTCGAGTTT	418	1263	CAGCAGCTTCCAACTATAATGTCTACTTTCAGTGAATGTTGACATTTGTTATTCCTCAGAC
QY	425	ATACCAACCATTTATGCGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGCACTGTG	484	1322	CGCTATGTTCCAACTTACACTCCCTCGTGTCTATCAGCTATTATCGTATCTGCAATGCTTG
Db	419	GTACCGCCATTTGTATAACGGTGTCTAGGAGTTTCAAGGACTTAGCGGTGGAACTG	478	1314	CGCTATGTTCCAACTTACACTCCCTCGTGTCTATCAGCTATTATCGTATCTGCAATGCTTG
QY	485	GCAGTTGGATCGCTTCTGATGGGTCGATGTTGAGTAAATGCCCTTGATCCCAATGAAGAC	544	1382	CGCTATGTTCCAACTTACACTCCCTCGTGTCTATCAGCTATTATCGTATCTGCAATGCTTG
Db	479	GCAGTTGGCTCTCTGTTGACAGGTGCAATGCTGAGCAAAAGTTCATGCTGAGAAAGAT	538	1374	CGCTATGTTCCAACTTACACTCCCTCGTGTCTATCAGCTATTATCGTATCTGCAATGCTTG
QY	545	CCAAAGCTTATACC-TCCACCTGGCTTTTCACAGCTACATTTATTTGCTGGTGTCTTCAGGC	603	1383	GACTCATAGATTATGAAGCAGCATCCATCTATTATTAAGTTCACAAATTTGACTTTGTGG
Db	539	CCTAAGCTTTACCAATTCACCTTCTTACCGCCACTTTTTCGCGCGGCTTCGGAAGC	598	1434	GACTCATGACTATCAAGCTGCCATCTCTGGAAGTTGACAAAGTTGCAAGTTCTCCTCG
QY	604	TGCTT-TGGCTCTGTTTAGTTCGGGTTCATCGTGGATTTCTCTGCATGCAACCATAA	662	1443	TGTGATGATGCTGATGCTGCTTCTTCTATTTATTTGCAAGGCCAAGCAATTCGTTT
Db	599	CTCATCTTGGAAATTTTCAGGTAGGTTTCATAGTGGACTTTCATCGCATGCAACGATAG	658	1502	CGTGTGCGATATCTTATAGCGAGTTTGTCTGTGTTGTCGAGGCCAAAACTGCGGTGA
QY	663	TAGGTTTCATGGAGAGCAGCCACCGTGTGTCTGACGAACTAAATTCGATTTCTTG	722	1563	TGGCAACATTTCCAAATCTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAAC
Db	659	TAGGATTCATGGAGGAGCAGCAACGCTGGTGAAGTCTGCAACAGATTAAGGGTATTTAG	718	1622	ACTTGGCTGAAAGAAATCATAGGTGGATGATGAAGAGGAGAGAGTAAACAATCAG
QY	723	GCCTTCAGCATTTTACCCATGGAGCTGATATCATATCAGTGTAGCTGTGTTTTCACCC	782	1743	GGGAGACTAGTTTTCAGTATGTTATAATTTGATATGAGTCTCTTGGAAACATTGTACAA
Db	719	GAATTCAAATTTTCACAGACTCAACAAATGTTATACAGTTCATGCAATCAATTTTTCACAA	778	1794	GAGAGAGCAGCTTACAATATATTTACTCGATATGTCAGTGTGTTGTAATATCGACAAA
QY	783	AAACTCATAGTGGAGTGGGAAGTGTCTGTGTAGGATGTGCTTCATTTTCTTCTCTCC	842	1803	GTGGAATAGTATGCTTGAAGAGGTGAAGAAATTTACAGAGAGAGAGAGCTACAGCTTG
Db	779	AAACACAAAGTGGAGATGGGAAGTGGCGTTTCATGGAATGGTTTTCATTTCAATTCATC	838	1795	GGGATATTAGCATGATGGTGGAAATTAAGAAATGATCAATGACAGGAGAGGTTAAAGTTGG
QY	843	TTAGCACAAGATCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGC	902	1863	TTTTGTCAATCTCTGTAGTGAAGTGAAGAAATGAAAGAAATGAAATTCGAAGT---TCCAAA
Db	839	ACACAACAAGATATTTTCAGCATCAAGCAACCAAAATTCATTTGGTGGGCGGATGGCTC	898	1855	TATTTGTCAATCCAAAAGGAGAGGTCGTGAAGAAATTAACCAAGATCCAAATTCATCGGTG
QY	903	CATTGACGTCGTTATATTGGGAAGTCTCTTGGTTTATTTTCACTCAGCGGAGAACG	962	1920	ATCATTTAGGGAAGAAATGGATCTATCTGACTGTTTGAAGAGCCGTTGAGGATGCAACT
Db	899	AATTGACCAAGTGTATTTTGGCAGTTCATATGTTTTCATCACAACACAGAGAGACATG	958	1915	ATCATTTGGGCAAGAGTGGATGTTCTTAACGTTAGGAGAAGCAGTGGAGGCTTGTAGCT
QY	963	GTGTTGAAGTATAGGAGAACTGAAGAAAGGTTTGAATCCACCATCACTCACAAAATCTGG	1022	1980	TCAATCTACGTCAAGCAAAACGAAACC 2007
Db	959	GTGTTCAAGTTCATAGGGCCAGAGACAGG---TTGCTTCACTCACAAGTTCTGTATCTCA	1015	1975	ACATGCTTTCACAGTTTAAACCGAACC 2002
QY	1023	TATTTGTGCTCCTTACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATCATAT	1082		

RESULT 2
CL961102
LOCUS
DEFINITION
ACCESSION

CL961102 1959 bp DNA linear GSS 21-SEP-2004
OeIFCC005483 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
CL961102

VERSION	CL961102.1	GI:52376926	
KEYWORDS	GSS.		
SOURCE	Oryza sativa (indica cultivar-group)		
ORGANISM	Oryza sativa (indica cultivar-group)		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE	1 (bases 1 to 1959)		
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.		
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.		
FEATURES	Location/Qualifiers		
source	1..1959 /organism="Oryza sativa (indica cultivar-group)" /mol_type="genomic DNA" /db_xref="taxon:39946" /clone_lib="Oryza sativa Express Library" /note="Oryza sativa exon trapped genomic sequences"		
ORIGIN			
Query Match	27.4%; Score 670.4; DB 9; Length 1959;		
Best Local Similarity	60.3%; Pred. No. 7.2e-164;		
Matches 1109; Conservative	0; Mismatches 731; Indels 0; Gaps 0;		
QY	126	ACCAAGTCGAGGTTCCACCCGACGCGTTTTCAGTCTCTAAAGTACTCTTTGGAAGG	185
DB	83	ACATGTCGGCGCGCCGCCAAGAGAAGAACTCTTCGGCGAGTTGCGCGCACGGTGAAG	142
QY	186	AGACTTTCTTCCCTGTAGTACCTTTTGAGGCGAGTTTCAAGAACAAAGCAGCTTCCAAGAAGT	245
DB	143	AGAGGTTCTTCTCGAGAGCGGATGCGCGGTACAGGACCGAGGTCAGGAAGC	202
QY	246	TCATGCTTGGCCTTCAGTCTTCTTCCCTCCATTTTCGAATGGGCTCCCAATAACACCTTTC	305
DB	203	TGTGGCTCGCCCTGCAGCAGTCTTCCCGGTGTTTGAATGGGCGAGCAATACACCCCTCG	262
QY	306	AGTTCCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAGG	365
DB	263	CCAAGTTCAAGGGCGACCTCATTTGCGGCTCCACCTTGGCAGCTCGTCATACCTCAGG	322
QY	366	GCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTTGAGCTATATTCGAGCTTTA	425
DB	323	ACATCGCTACGGGAAGCTTGCTAACTTGCCACAGAGATTTGGCTGCCAAGTAGCTTCG	382
QY	426	TACACCAATTTATTCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGAGCTGTGG	485
DB	383	TCCCGCCATTTGATATACGCTCTGATGGGCACCTCAAGGAGCTAGCAATGGGTCCAGTGG	442
QY	486	CGGTTGGATCGCTTCTGATGGGTTCGATGTTGAGTATGCGGTTGATCCCAATGAAGACC	545
DB	443	CCGTCATCTCACTGCTCTGTTGTAATCTCTCCAGGAGGAGATTTGACTCAAGAAGAAC	502
QY	546	CAAAAGCTTACCTCCAGCTTTCACAGCTACATTTATTTGCTGGTGTGTTTTCAGGCTG	605
DB	503	CGCTAGATTACAGAGGCTCGCCTTCACAGGACCTTCTTTGACGGAGTCCACAGGCGG	562
QY	606	CTTGGGCTGTTTATGATGATGGGTTGATCGTGGATTTTCTGTCTCAGTGCACCAACATATAG	665
DB	563	CGCTTGGTTCCTGAGGCTAGGGTTTCATAGCGTTCCCTGCTCATGCTGCCATCATCG	622
QY	666	GGTTCATGGGAGNGAGCCACGGTGGTGTGTCTGCGAGCAACTAAATTCGATCTTCTGGCC	725
DB	623	GATTCATGGCGCGCGCCCATCACCATTGCTCTTTCAGCAGCTTAAAGGCTTCTTGGA	682
QY	726	TTGAGCATTTTCAACCATGGAGCTGATATCATATCAGTGATCGCTCTGTGTTTCAACCAAA	785
DB	683	TTGCAAAATTTCAACCAAGAGACTGACATCATCTCGGTTCATGAAATCAGTCTGGGAATG	742
QY	786	CTCATGATGGAGGTGGGAAAGTGTGTGTAGATGTGTCTTCATTTTCTTCTCTCTTA	845
DB	743	TTCCACCATGGGCAATGGAACTGCGCAGACAATATTGATCGGAGCATCATTTTGGCATTC	802
QY	846	GCAACAAGATACTTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTACAGCAATGCGCCAT	905
DB	803	CTCTGTTGTCGAAGGCAAGAGAAACAAGAGCTCTTCTGGTCCAGCAATTTGCAACAC	862
QY	906	TGACGTCGCTTATATTGGGAAGTCTCTTGGTATTATTTCACTCACGCGGAGAACCGGTG	965
DB	863	TCATTTTCGGTGATCATTTTCAACTTTGTTGCTCTACATCACTCGTGTGACAAAACAGGCG	922
QY	966	TTGAAGTGATAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATCTGGTAT	1025
DB	923	TCGCAATTTGTAATAAAGCGTCAAGAAAGGCATCAATCCACTTCAGCTAGCTGATATTTT	982
QY	1026	TTGTGTCGCTTACATGACTACAGCTGTCAAACTGCGCTTGTGCTTGGCATCATATCAC	1085
DB	983	TCATGCGCCATCTTCTCAAAGGATTCAAAATTTGAGTAGTAGCTGGAATGATAAGCC	1042
QY	1086	TTGCGGAAGGAATAGCAGTAGGAGAAAGCTTTTCAATGTATATAAAATTTACAATATTGATG	1145
DB	1043	TTACGGAAGCGATTGTCAGTTGGAAGAACAATTTCTGGATTGAACGATTACCAGATAGATG	1102
QY	1146	GCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTCACTCTTGCT	1205
DB	1103	GGAAACAAAGAAATTTGGCTCTAGGAAACCATGAATGTGGTGTGCTCAATACGCTTTGCT	1162
QY	1206	ACCTCAACAGGACCATTTTTCGCTGCTGAACATAACGCTGGAATGCAAGACAG	1265
DB	1163	ATATAGCCACAGTGGTTTTCAGCATCAGAGTCAATTCGATGGCTGGAGGTAAACAC	1222
QY	1266	CAGCTTCCAACATTAATATGTCATTTGCAGTAATTTTGAATTTGTTATTTCTGACACCT	1325
DB	1223	CAATGTCCAATATTGTTATGTCACTGATGATTTGCTTGCACTCTCTGTGGATCACTCCAT	1282
QY	1326	TGTTTCATTTACACTCCCTGGTGTGCTATCAGCTATTATCGTATCTGCAATCGTGGAC	1385
DB	1283	TGTTCAAGTACACGCCCAATGCCACCAATTTCTTCATCATCATATCAGCAGTCTTGCC	1342
QY	1386	TCATAGATTATGAAGCAGCCATCCATCTATTAAAGTTGACAAATTTGACTTTGTGTGT	1445
DB	1343	TATTTGACTTTGAAATCAGCTACCTTATCTGGAAGGTTGATAAGTTGGACTTCATGCGGT	1402
QY	1446	GCAATGATGATACATTTGGCGTGTCTTTGGCAGTGTGGAATTTGGCTTAGTCAAGCTA	1505
DB	1403	GCTTGGGGCATTTCTTTGGAGTAATATTTTCATCTGTGGAGTATGGCTTCTCATTCGGG	1462
QY	1506	TTGTAATATCTGACTCTGGGTACTTCTATTTATTCGAAGGCCAAGACATTCGTTTGG	1565
DB	1463	TTGTAATATCACTTAATCAAAAGTTCTGCTCCATGTAACACGGCCAAAGCAGCTTTACTTG	1522
QY	1566	GCAACATTTCAAAATCTGTGATATACGAAATGTTGAGCACTATCAAAATGCAAAACATG	1625
DB	1523	GCAACCTTCCAGACGATTTCTATAGAAATTTGAAACAAATATCCAGAACTACCAAGG	1582
QY	1626	TTCTCTGGAATGCTTAATTTCTAGAGATTGATGCACAATTTTACTTTTGCAATGCCAGCTATT	1685
DB	1583	TGCCAGGATGCTTAATTTGAAGAGTGGACTCAGCAATATACTTCAACAACTCCAATATG	1642
QY	1686	TAAGAAAGGATCAACAGGTGATTTGATGAAGAGAAAGAAATTTAAAGCTACAGGG	1745
DB	1643	TTAAAGAAAGAAATGCTGAGATGGCTGAGAGATGAGGAAGAACATCAAAAGGAAACAGAGT	1702
QY	1746	AGACTAGTTGCAAGTATGTTATTAATTTGATATGATGCTGTTGAAACATTTGATACAGTG	1805

```
Db      1703  TACCAAAAAATTGAGTTTCTGATTGTTGACCTATCTCCTGTAATGATATTGACACAAGTG 1762
Qy      1806  GAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGTACAGCTTGTTT 1865
Db      1763  GAATCATGCTTCAAGAGATTGTTGAGGACATTGAAGGCCAGATTACGCTGATTT 1822
Qy      1866  TGGTCAATCCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTTCCAAAATCAT 1925
Db      1823  TCGCCAATCCTGGGGCGGTGTGATCCAAAAGCTCCGGTCAGCAAAATTCACAGAGCTCA 1882
Qy      1926  TAGGGAAGAAATGAGTCTATCTGACTGTTGCAAGAGGCCGT 1965
Db      1883  TTGGTGAAGAAAAAGATATGCTGACAGTTGGTGACGCCGT 1922

RESULT 3
LOCUS   CB892638
DEFINITION EST645430 HOGA Medicago truncatula cDNA clone HOGA-19G10, mRNA
ACCESSION CB892638
VERSION   CB892638.1
KEYWORDS  CB892638.1 GI:30099806
SOURCE   Medicago truncatula (barrel medic)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 854)
AUTHORS  Hahn,M.G., Ojanen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
Utterback,T., Cho,J. and Fraser,C.M.
TITLE    ESTs from roots of Medicago truncatula treated with
oligolacturonides of DP 6-20
JOURNAL  Unpublished (2001)
COMMENT  Contact: Michael G. Hahn
Complex Carbohydrate Research Center
University of Georgia
220 Riverbend Road, Athens, GA 30602-4712, USA
Tel: 706-542-4457
Fax: 706-542-4412
Email: hahn@ccrc.uga.edu
TIGR sequence name: MTMCV417K
More information is available at: www.medicago.org
Seq primer: Sknod (CTA GAA CTA gtg gat CC).

FEATURES             source
source               1..854
                    /organism="Medicago truncatula"
                    /mol_type="mRNA"
                    /cultivar="A17"
                    /db_xref="taxon:3880"
                    /clone="HOGA-19G10"
                    /tissue_type="3 day old seedling roots"
                    /dev_stage="24 hours after treatment in the dark at 26 C
                    with 0.5 mg/ml oligolacturonides (DP 6-20) in the
                    presence of 100 ug/ml Gentamicin"
                    /lab_host="XL0LR"
                    /clone_lib="HOGA"
                    /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
                    XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                    was directionally ligated into the Unizap XR vector from
                    Stratagene and packaged using Gigapack III Gold packaging
                    extracts. Plasmids containing cDNA inserts were excised
                    from the recombinant lambda-Zap phage using Ex-assist
                    helper phage and propagated in SOLR cells."

ORIGIN
Query Match      26.8%; Score 655.6; DB 6; Length 854;
Best Local Similarity 85.5%; Pred. No. 4.2e-160;
Matches 730; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy      273  CGATTTTCGATGGGCTCCAAATACACCTTTCAGTTCTTGAAAGCTGACCTCATAGCTG 332
|||||
```

```
Db      1    CCATATTTGAATGGGCACCAAGCTACACTTTTTCACCTTTTGAATCTGATCTCATAGCTG 60
Qy      333  GCATCACCATCGCTAGCTTGGCCATTTCCTCAGGGCATCAGTTTATGCCAAGCTCGCCAAAC 392
|||||
Db      61  GCATCATAATTCCTAGCTTGGCTATTTCCTCAAGGCATTTAGTTATGCCAAACTTGGCAAC 120
|||||
Qy      393  TCCCTCCAATTTCTGGACTTATATTCGAGCTTTATACCAACCATTTGATTTATGCGATGATGG 452
|||||
Db      121  TCCCTCCAATTTCTGGACTTATATTCAGGCTTTATACCAACCGTTAAATTTATGCAATGATGG 180
|||||
Qy      453  GTAGCTCGAGGAGATTGGCAGTGGGACATGTGGCGGTGGATCGCTTCTGATGGGTTTCA 512
|||||
Db      181  GTAGTTCAAGGAGATTGGCAGTTGGAATGTGGCAGTTGGATCTCTTCTGATGGGTTTCA 240
|||||
Qy      513  TGTGTAGTAAATGCCGTTGATCCCAATGAAGACCCAAAGCTTTTACCTCCACCTGGCTTTCA 572
|||||
Db      241  TGTGTGCCAATGAGGTTTAAATCTTACCCAAAATCCAAAGCTTTTCTCTGCACTTTCCTTTCA 300
|||||
Qy      573  CAGCTACATATTTCGTGGTGTGTTTTTCAGGCTGCCTTGGGTCCTGTTAGTTTGGGTTGA 632
|||||
Db      301  CAGCTACATTTCTTGTGGACTTTTTCGACGCTTCATTTGGGTCCTCTTTAGGTTAGGGTTTA 360
|||||
Qy      633  TCGTGGATTTTCTGTCACATGCAACCATTAATAGGGTTTCATGGGAGAGCAGCACCGTGG 692
|||||
Db      361  TCGTGGATTTCTGTCACATGCGCCATAGTAGGGTTTCATGGGAGAGCAGCACCGTGG 420
|||||
Qy      693  TGTGCTCGAGCAACTAAATCGATTCTTTGGCCTTGGAGCATTTTCAACCATGGAGCTGATA 752
|||||
Db      421  TTTGCTACAGCAGCTGAAATCAATCTTAGGACTTTGAACATTTTACCCCATGCTGCTGATA 480
|||||
Qy      753  TCATATCAGTATGCGCTCTGTTTTTACCACCAACTCATGAGTGGAGTGGGAAAGTCTCT 812
|||||
Db      481  TTGTATCCGTCATGCGCTCCGTTTTTACCACCAACTCATCAGTGGAGTGGGAAAGCGCTG 540
|||||
Qy      813  TGTTAGGATGTCCTTCATTTTCTTCTCTTAGCACAAGATACCTTCAGCAAAAACGAC 872
|||||
Db      541  TGTGGGATTTTGGTTCATTTTCTTCTCTTTTACAAAGATACCTTCAGCAAAAACGAC 600
|||||
Qy      873  CAAAGTTTTTTTTGGGTGTCAGCAATCGGCCCATTCGCGTCCGTATATTTGGGAAAGTCTCT 932
|||||
Db      601  CCAAGTTCTTTTGGGTGTCAGCAATGACGCCGTTAGCGTCCGTTATATCTGGCGAGTTTGC 660
|||||
Qy      933  TCGTTTATTTTCACTCAGCCCGAGAGACGCGGTGTTGAAGTATAGAGAACTGAAGAAAG 992
|||||
Db      661  TGGTTTATTTCACTCATGCGGAGCACCAGCGGTTCAAGTATAGAGAAATTTGAAGAAAG 720
|||||
Qy      993  GTTTGAATCCACCATCACTCACAATCTGGTATTTGTGTCGCTTACATGACTACAGCTG 1052
|||||
Db      721  GACTAAATCCACCATCACTCAGAGATTTGGTATTTTGTGTACCATATATGACACACGTA 780
|||||
Qy      1053  TCAAAACTGGCATTTGCTTGGCATCATATCACTTGGCGAAGGAATAGCAGTAGGAAGAA 1112
|||||
Db      781  TCAAAACAGGCCCTTATTTGTCATTTATAGCTCTTTGGGGAAGGAATCGCTGTGGGAGAA 840
|||||
Qy      1113  GCTTTGCAATGAT 1126
|||||
Db      841  GCTTTGCAATGAT 854
|||||

RESULT 4
LOCUS   AY105934
DEFINITION Zea mays PCO089998 mRNA sequence.
ACCESSION AY105934
VERSION   AY105934.1
KEYWORDS  AY105934.1 GI:21209012
SOURCE   Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 2279)
AUTHORS  Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
```

Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2279)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source
1..2279
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635109"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 26.7%; Score 653.8; DB 3; Length 2279;
Best Local Similarity 60.3%; Pred. No. 1,7e-159;
Matches 1119; Conservative 0; Mismatches 727; Indels 9; Gaps 2;
QY 121 AGTGCACCAAGTCAGGTTCCACGCCACACGCGTTTTTCAAGTCTCTAAGTACTCTTT 180
DB 188 AGTGCAACAGGTGGCGCGCCACCGCGCGGAGCAGCGGAGCAAGATGAAGTGAGGGT 247
QY 181 GAAGGAGACTTTCTTCCTGATGACACCTTTGAGCAGTTCAAGAACAAAGCAGCTTCCAA 240
DB 248 GAAGGAGACCTTTCTCCCGACGACCGTTCCGGCGTTCAAGGGCAGCGCGGGGAC 307
QY 241 GAAGTTCAAGCTTGGCTTCAGTTCTTTCCCAATTTTCGAATGGCTCCCAAAATACAC 300
DB 308 GCAGTGGCTCATGCGCGTCAAGTACCTCTTCCCATCTCTGACTGGGTGCGAGCTACTC 367
QY 301 CTTTCAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCC 360
DB 368 CTTTGCCTCTTCAAGTCCGACCTCGTCGCGGGCTCACCAATGGCAGCCTCGCCATTCC 427
QY 361 TCAGGGCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTGGACTATATTCGAG 420
DB 428 TCAGGGCATTTAGCTACGCGAAGCTGGCAAGCTTGGCTCCCATTAATCGGCTGTATTCGAG 487
QY 421 CTTTATACCAACCATGATTATGCGATGATGGGTAGCTCAGGGATTTGGCAGTGGGAC 480
DB 488 CTTTCGTGCGCGCATGGTGTACGCGGTCTGGGAGCTCCCGTGACCTGGCGGTGGGCC 547
QY 481 TGTGGCGGTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTGATCCCAATGA 540
DB 548 GGTGTGATCTGTCGCTGATCATGGGGTCCATGCTGCGGACGCGCGTGGCCCAACCGC 607
QY 541 AGACCCAAAGCTTTACTCTCACCTGGCTTTACAGCTTACATTAATTTGCTGGTGTTTTCA 600
DB 608 GGAGCCGACGCTGTCTGCGAGCTGGCTTCACTCACCTGTTGGCGGGCTGGTGCA 667
QY 601 GGCTGCTTGGGTCTGTTTGGGTGGGTGATCGTGGATTTTCTGTCAATCAACCAT 660
DB 668 GGCCTCCCTGGGCATCTCAGGCTCGGCTTCGTTCATCGACTTCTGTGCCAAGCGCGCT 727
QY 661 AATAGGGTTCATGGGAGGACGCCACGGGTGGTGTCTGCGCAACATAAATCGATTCT 720

DB 728 GGTGGGGTTCATGGCCGGCGCCCATCATCTGTTGGCGCTGCAGCAGTCAAGGCGTGTCT 787
QY 721 TGGCCTTTGAGCAATTTCCACCATGGAGCTGATATCATATCAGTGATGGCTCTCTTTTTCAC 780
DB 788 GGGCATCGTCCATCTTCCACCACGAGATGGCATCGTCCAGTCATGGCTCTTCCCTTCCA 847
QY 781 CMAAATCATAGTGGAGGTGGAAAGTGTGTGTAGGATGTCTTCAATTTTCTTCTCT 840
DB 848 CCACACCAGCGATGGTCTGGGCGAGACGATCTCTCATGGGCGTCTGCTCTCTCTCTCT 907
QY 841 CCTTAGCAACAAGTACTTCCAGCAAAAACGACCAAGGTTTTTTTGGGTGTGCACAATGGC 900
DB 908 GCTGTCCGGAGGCAATGTGAGCATCAGATGGCCAAAGCTTTTCTGGGTTTGGCGGTGCGC 967
QY 901 GCCATTGAGTCCGTTTATATTGGGAAAGTCTCTTTGGTTTATTTTTCACCTCACGCCGAGAGCA 960
DB 968 GCCCTGGCATCGGTCAACCATCTCGACGCTGCTTTTCTCTTCAAGCTCAGAACCA 1027
QY 961 CGGTGTTGAAGTATAGGAGAACTGAAGAGGGTTTTGAATCCACCATCACTCAAAATCT 1020
DB 1028 TGGCATCAGCATCATTTGGGCGAGCTCAAGTGGCGCTGAATCGCCCTCGTGGGACAAGCT 1087
QY 1021 GGTATTGTGTGCGCTTACATGACTACAGCTGTCAAACTGGCATGTGTTGGCATCAT 1080
DB 1088 CTTGTTTGACACGGCGTATTTAGGCTCACATGAAGACTGGGCTTGTCAACGGAATCAT 1147
QY 1081 ATCAGCTTGGGGAAGGAATAGCAGTAGGAAGAAGCTTTTGCAAATGTATAAAATTTACAAAT 1140
DB 1148 CTCAGCTGACGGAAGGAATAGCGGTTGGTAGACATTTGCTCTCACTCAAGGACTACCGAT 1207
QY 1141 TGATGCCAACAAAGAGATGATAGCTATTTGGGACCATGAACGTTAGTTGGTTCTTTTCACTC 1200
DB 1208 AGATGAAACAAAGGAGATGATGGCCATAGGGTTGATGAATGTTTGGGTCTCTGCACATC 1267
QY 1201 TTGCTACCTCACAAAGGACCACTTTTCGCGTTCGCGTGTGAACATAAAGCTCGATGCA 1260
DB 1268 ATGCTACGTAAACAAGGTGGCTTCTCCGCTCTGCTGTAACCAACAGCCGCGGTGCA 1327
QY 1261 GACAGCAGCTTCCAAACATTAATTAATGCTCACTTGCAGTAATTTTGACATTTGTTTCTCTGAC 1320
DB 1328 GACTGCCATGTCCAAGCTGATCATGCGCTGACTGTGATGGTCACGCTGCTGTCTCTCAT 1387
QY 1321 ACCCTTGTTCATATACATCCCTGGTGGTGTATCAGCTATTTATCGTATCTGCAATGCT 1380
DB 1388 GCCACTGTTCTGTGTACACACCAACGTTGCTCTCGGAGCGATCATCATCGCCGCGTGT 1447
QY 1381 TGGACTCATAGATTATGAAGAGCCCATCCATCTATTTTAAAGTTGACAAATTTGACTTTGT 1440
DB 1448 CGGCTGTGATCGATTTCCCGCGGTGTACCAATCTGGAAGATGGACAAGATGATTTCT 1507
QY 1441 GGTGTGCATGAGTGCAATATTGGCGTGTCTTTTGGCAGTGTGAAATTTGGCTTAGTTCAT 1500
DB 1508 GGTGTGCGTTTGGCGTGTTCGCGCTCATCTTCTCATCTCAGTCCAAGAGGCTTGGAT 1567
QY 1501 AGCTATGTAATATCTGTACTCTTGGGTACTTCTATTTTATTTGCAAGGCCAAGGACTTCGT 1560
DB 1568 AGCGTGTGTATATCTATATTTAGGGTGTGATGAGATCAACAAGGCCAAGATGATGTGT 1627
QY 1561 TTTGGGCAACATTTCCAAATTTCTGTGATATACCGAAATTTGAGCACTATCAAAATGCAA 1620
DB 1628 TCAAGGGAACATCAAGGGGACTGATATTTACAGAGACCTGTCATCTACAAAGGAGGCCA 1687
QY 1621 ACATGTTCTCGGAATGCTAAATTTAGAGATTTGATGACCAAAATTTTACTTTGCCAATGCCAG 1680
DB 1688 AAGAGTTTCTGGTCTTCTGATCTTGGCCATTTGAAGCACCGATAAACTTCCGCAACTCCAA 1747
QY 1681 CTATTTAAGAGAAAGATCAAGGTGGATTTGATGAAGAAGAAAGAAATTTAAAGCTAC 1740
DB 1748 CTACCTGAAATGAAGGATTTAAAGATGGAT-----AGAGGAAGAAATCTTTTGAACAGGA 1801
QY 1741 AGGGGAGACTAGTTTGCAGTATGTTTATAATTTGATGATGCTGTGTTGGAAACATTCATAC 1800

Db	1802	TAAACATACTGAACCTCAATTCATTAATCTTGGATCTGTCTGAGCTGTCTTCTGCAATTCGACAC	1861
QY	1801	AAGTGGAAATAGATATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGCTACAGCT	1860
Db	1862	AAGTGGCATAGCGTTCTCTATTGACATTAAGAAATCAATAGAAACGGTCTGAGCT	1921
QY	1861	TGTTTTGGTCAATCTCTGTAGTGAAGTGATGAAGAACTGAAACAAATCGA--AGTTCCA	1917
Db	1922	TGTGCTTGTCATCACTCAACTGGAGAGTCAATGGAGAAATACACGTGCAACGAGGCTGA	1981
QY	1918	AAATCATTTAGGAAGAAATAGATCTATCTGACTGTGTGAAGAGCGCGTTGGAGCA	1972
Db	1982	AAACTATTTTAGGCCAGATTGCTTGATCTGACCACTGGCGAAGCAATCGCTTCA	2036
RESULT 5			
CL959635			
LOCUS			
DEFINITION	CL959635	Oryza sativa (indica cultivar-group)	
ACCESSION	CL959635	Oryza sativa (indica cultivar-group)	
VERSION	CL959635.1	GI:52373944	
KEYWORDS	GSS.		
SOURCE	Oryza sativa (indica cultivar-group)		
ORGANISM	Oryza sativa (indica cultivar-group)		
REFERENCE			
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.		
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.		
FEATURES			
source	1. .2001		
	/organism="Oryza sativa (indica cultivar-group)"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:39946"		
	/clone_lib="Oryza sativa Express Library"		
	/notes="Oryza sativa exon trapped genomic sequences"		
ORIGIN			
QY	145	GCACAGCGCTTTTCAAGTCTCTAAGTACTCTTTGAAGGAGACTTCTCCCTGATGA	204
Db	123	GGCCGCCCGTTCGGCGGAAGCTGTGTCGGACCTCGCGGAGACGTTCTTCCCGACGA	182
QY	205	CCCTTTGAGCGAGTTTCAAGAAACAGCCAGCTTCCAAAGAAAGTTTCATCTTGGCTTCAGTT	264
Db	183	CCCGTTCCGGGGTTCGGCGCGTCCCGCGCCAGCGCGGTGGTGGCGGTCAAGTA	242
QY	265	CTTCTTCCCAATTTTCAATGGGCTCCCAAAATACACTTTTCAAGTTCTTGAAGAGTGACCT	324
Db	243	CTTCTGTCGGCGGCTGGAGCTGGGTCCTCGGCTACGGCTCGAAGTTTCAAGTTTCAGCT	302
QY	325	CATAGCTGGCATCACCATCGCTAGCTTGGCCATTCTCAGGCGCATCAGTTATGCCAAGCT	384
Db	303	CCTCGCGGCGATCACCATCGCCAGCTTCGCCATCCCGAGGGGATCAGCTACGCCCGCT	362

QY	385	CGCCCAACCTCCCTCCAATTTCTTGGACTATATTCGAGCTTTATACCAACCAATTCATTTATGC	444
Db	363	CGCCCAACCTCCCTCCAATTCATCGGCTCTATTCGAGCTTCGTCGCCCGTTCATGTACGC	422
QY	445	GATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTGTGGCGTTGGATTCGTTCTGAT	504
Db	423	CGTGTTCGGGAGCTCCAAACAACCTGGCGTGGGACGGTGGCGGGCGGCTCGTGTGCT	482
QY	505	GGTTTCGATGTTGAGTGAATGCCGTTGATCCCAATGAAGACCCAAAGCTTTTACCTCCACCT	564
Db	483	GGCGTCCATCATCGAGACGGAGGTGGCGGCGACAGAAACCCGAGCTGTACTCTGAGCT	542
QY	565	GGCTTTTCACAGCTACATTTATTTGCTGGTGTGTTTTTTCAGGCTGCTTTGGGTCTCTTTAGGTT	624
Db	543	CTTCTACACCGCGCTTCTTCAACCGGCTCTTCCAGACCGGCTAGGCGCTTTTAGGTT	602
QY	625	GGGGTTGATCGTGGATTTTCTGTCAATGCAACCAATTAATAGGGTTTCATGGGAGGAGCAGC	684
Db	603	AGGGCTAATAGTGGATTTTCTCTCGCGGTTCGACGATCACCGGTTTCATGGGCGGACGCGC	662
QY	685	CACGGTGGTGTCTCTGCAGCAACTAAATTCGATTCTTGGCCTTGAGCATTTTCACCCATGG	744
Db	663	GATGATCATCATCTACTGACAGCTTCAAGGCGCTCTCTGGGATGAAGACCTTCACCAACAA	722
QY	745	AGCTGATATCATATCAGTGAATCGCTCTGTGTTTTTACCCCAAACTCATGAGTGGAGGTGGA	804
Db	723	GACAGACATCATCTCGTCTCTCACTCCACCTACCATTACCGGACGAGTGGAAATGGCA	782
QY	805	AAGTGTGTTAGGATGTCTTCAATTTTCTCTCTCTAGCACAAAGATCTTCAGCAA	864
Db	783	GAGTGCAGTTCTCGGCATATGCTCTCTCTTGTCTGTCTGCTGAGTAAGACCTTCAGAAA	842
QY	865	AAAAACGCAAGGTTTTTTTGGGTGTCAGCAATGCGCCATTGAGCTCCGTTATATTGG	924
Db	843	GAATGCGCAAGGTTGTTCTGGGTGTGCGCCATTGCGCCATTTCATGTTGTGCTCATCG	902
QY	925	AAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGACGCTGTGTAAGTGATAGGAGAACT	984
Db	903	AGGCATCTTTCGCTTCTCTCGTCAAGGGTGACGAGCATGGAAATTCGGATAGTCGTAACCT	962
QY	985	GAAGAAGGTTTGAATCCACCATCACTCACAATCTGGTATTTGTGTCGCCCTTACATGAC	1044
Db	963	GAATAAAGGGATCAATCCCTCTGTCATTTGCGCAGCTAACGTTTCCAGTCAGGACATGAA	1022
QY	1045	TACAGCTGTCAAACTGCGATTGTCTTGGCATCATATCACTTGGGAGGAATAGCAGT	1104
Db	1023	GACGCGAGTGAAGCTGTGCTCTTCTGCGGATCTTAGCAGTGGCAGAGGAATAGCCGT	1082
QY	1105	AGGAAGAAGCTTTGCAATGTATPAAAAATTAACAATATTGATGGCAAACAAAGAGATGATAGC	1164
Db	1083	CGGCCGAAGCTTGGCGATGGTCAAGAACGAGCAGATCGACGGCAAACAGAGAGATGATCGC	1142
QY	1165	TATTGGGACCATGAACGTAGTTGGTTCTTTACCTCTTGTCTACCTTCACAAAGGACCATTT	1224
Db	1143	GTTTGGTATCATGAACATCGCGCGTTTCTTTCACCTCTCTGCTACCTCAGCACAGGCGCGTT	1202
QY	1225	TTCCGCTTCGGCTGTAACATAAAGCTGGATGAAGACAGCAGCTTCCAAACATTATATAT	1284
Db	1203	CTCAAGTTCGGCGGTGAACCTCCACGCCGGGTGCAAGACGCCGATGTGCAAGCTGGTAT	1262
QY	1285	GTCACTTCGAGTAATGTTGACATTTGTAATTCCTGACACCCCTTGTTCATTTACATCCCT	1344
Db	1263	GTCCGTGTGCATCATGCTGGTGTACTGTCTTCTGGCACCCCTCTTCAAGTACACCCGCT	1322
QY	1345	GGTGTGTATCAGCTATTTATGCTATCTGCAATGCTTGGACATCATAGATTATGAAGCAGC	1404
Db	1323	GGTGGCGCTCTCTCCATCATCTGTCGCCATGATCGGGCTCGTCAAGGTCAGAGGTT	1382
QY	1405	CATTCATCTATTTAAGGTTGACAAATTTGACTTTTGTGGTGTGATGAGTGCATATATTCG	1464
Db	1383	CTGCCACCTCTACAAGTTCGACAGTTGACTTTCTGCACTGTCATGGTTCGCTTCTCTCG	1442
QY	1465	CGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCTAGCTATTTGTAATATCTGTACTTCG	1524

1443 CGTGTCTTCTTACCAATGTCGAGGCTCAGGCGATCGTAGGCTTGCAAGTTGTCAG 1502
1525 GGTAATCTTATTTATGCAAGGCCAAGGACATTCGTTTTGGGCAACATTCACAAATTCGT 1584
1503 GGCCTGCTGTACGTGCGCAGGCTGCACCTGCAAGCTCGGGAACATAGCAGGAGCGA 1562
1585 GATATACGAAATGTTGAGGACATCAAAATGCAAAACATGTTCTCGAATGCTAATCT 1644
1563 GACCTCCGCGACGTGAAGCAATACCCCAAGCGAAGAGCATCCCTGGCATCTCGCTT 1622
1645 AGAGATTGATGACCAATTTACTTTGCAATGCCAGCTATTTAAGAGAAAGGATCAACAG 1704
1623 CGAGCTGGCTCTCCATCTACTTCGTCATGCGGTTACTGCGAAGAAAGATTTTGAG 1682
1705 GTGGATTGATGAAGAAGAAAGAAATTAAGCTTACAGGGGAGACTGATTTTGCAGTATGT 1764
1683 ATGG--GTGGAAGATGAGGATACCTGTGCAAGAGCGTGGCGCAGCATTTGCAATACTT 1739
1765 TATTAATGATGATGCTGCTGTTGGAACATGTATCAAGTGGATTAAGTATGCTTGAGA 1824
1740 GGTCTTGTGATCTGGTGTGCTCACTTCTGTCGAACTCTGGAGTTGGGATGCTACTAGA 1799
1825 GGTGAAGAAGATTACAGAGAGAGAGACTACAGCTTGTGTTTGTCAATCTGTAAGTGA 1884
1800 AGTACACAAGAGCTCGAAAGAGAGGATCACGATAGCTCTGACGAATCCGAGGTAGA 1859
1885 AGTGATGAAGAACTGAACAAATC--GAAATTCCTCAAAATCAATTTAGGGAAGAAATGGAT 1941
1860 GGTGACAGAGAGCTGTGCTGTGATACGTCAGGAGCATCTTAGGGGATGAGTGGGT 1919
1942 CTATCTGACTGTGAAGAGCGCGTTGGAGATGCAACTCAATCTACGTCAAGCAAAAC 2001
1920 CTTCTGACGGTCAAGAGCGCATACGCGGTGTCGATACGCGTGCAGATATCCAGAA 1979
2002 GAACCCAAAGAGATGAA 2020
1980 TAAGGGAAGACGAGTA 1998

RESULT 6
AW598447
LOCUS
DEFINITION
BJ91904.y1 Gm-cl023 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl023-1999 5' similar to TR:064435 064435 SULFATE TRANSPORTER.
[2] TR:065025 ; mRNA sequence.

AW598447
AW598447.1 GI:7285969

EST.
Glycine max (soybean)

ORGANISM
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

1 (bases 1 to 583)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, B., Kohn, S., Shin, I., Jackson, F., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)

Other ESTs: BUS44910 corresponding to Gm-r1088-1045 (3')
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project

Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this

clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1570 Std Error: 0.00
High quality sequence stop: 440.

FEATURES

source

1. 583
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="T157"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl023-1999"
/tissue_type="seed coats of greenhouse grown plants"
/lab_host="DH108"
/clone_lib="Gm-cl023"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from seed
coats (100-200mg) of greenhouse grown plants. The
library was prepared using the Life Technologies
pSuperScript cDNA library construction kit. Complementary
DNA was synthesized from mRNA using a poly (dT) sequence
with a Not I restriction site. Sal I linkers adapters
were ligated to the blunt-ended cDNA fragments followed by
Not I digestion. The cDNA fragments were directionally
cloned into the Not I-Sal I restriction site of the
pSPORT1 vector. The ligated cDNA fragments were
transformed into E.coli Electromax DH10B host cells. This
library was constructed by Dr. Lila Vodkin and Dr. Anu
Khanna."

ORIGIN

Query Match 23.7%; Score 581.4; DB 2; Length 583;
Best Local Similarity 99.8%; Pred. No. 1e-140;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1198 CTCCTGCTACCTCACACAGGACCATTTTCGCGTTCGCGTGTGCACTTATTAATGTCACATTTGTCAGTAAATGTTGACATTTGTTATTCCT 1317
Db 1 CTCCTGCTACCTCACACAGGACCATTTTCGCGTTCGCGTGTGCACTTATTAATGTCACATTTGTCAGTAAATGTTGACATTTGTTATTCCT 120
Qy 1258 CAAGCAGCAGCTTCCCAACATTAATGTCACATTTGTCAGTAAATGTTGACATTTGTTATTCCT 1317
Db 61 CAAGCAGCAGCTTCCCAACATTAATGTCACATTTGTCAGTAAATGTTGACATTTGTTATTCCT 120
Qy 1318 GACACCTTGTTCATTAATGTCACATTTGTCAGTAAATGTTGACATTTGTTATTCCT 1377
Db 121 GACACCTTGTTCATTAATGTCACATTTGTCAGTAAATGTTGACATTTGTTATTCCT 180
Qy 1378 GCTTGACATCATAGATTATGAGCAGCCATCCATCTATTAAAGTTTGACAAATTTGACATT 1437
Db 181 GCTTGACATCATAGATTATGAGCAGCCATCCATCTATTAAAGTTTGACAAATTTGACATT 240
Qy 1438 TGTGCTGTGCATGAGTGCATACATTTGCGGTGCTTTTGGCAGTGTTCGAAATTCGCTTAGT 1497
Db 241 TGTGCTGTGCATGAGTGCATACATTTGCGGTGCTTTTGGCAGTGTTCGAAATTCGCTTAGT 300
Qy 1498 CATAGCTATTGTAATATCTGTACTTCGCGTACTTCTTCTTATTTATTCGAGGCAAGGACATT 1557
Db 301 CATAGCTATTGTAATATCTGTACTTCGCGTACTTCTTCTTATTTATTCGAGGCAAGGACATT 360
Qy 1558 CGTTTTGGGCAACATTTCCAAATCTCTGATATACCGAAATGTTGAGCAGCTATCAAAATGC 1617
Db 361 CGTTTTGGGCAACATTTCCAAATCTCTGATATACCGAAATGTTGAGCAGCTATCAAAATGC 420
Qy 1618 AAAACATGTTCTCGGAATGCTAATCTAGAGATTGATGACCAATTTTACTTTGCCAATGC 1677
Db 421 AAAACATGTTCTCGGAATGCTAATCTAGAGATTGATGACCAATTTTACTTTGCCAATGC 480
Qy 1678 CAGCTATTTAAGAGAAAGGATCAACAGGTGGATTTGATGAAGAGAGAAAGAAATTAAGC 1737
Db 481 CAGCTATTTAAGAGAGAAAGATCAACAGGTGGATTTGATGAAGAGAGAAAGAAATTAAGC 540
Qy 1738 TACAGGGGAGACTAGTTTCGAGTATCTTATTTATTTGATATGAGT 1780

Db	541	TACAGGGGAGACTAGTTTGCAGTATGTTATTAATGATAGT	583	QY	784	AACTCATGATGGAGTGGAAAGTCTGTGTTAGGATGTGCTTCATTTTCTTCCTCT	843
RESULT 7				Db	594	CACCAAGGAGTGGTCTGTCGACACGATCCTGATGCCGCTGCTTCCTGCTCTCTGCT	853
CL969864				QY	844	TAGCAAGAATCTTCAGCAAAAAACGACCAAGGTTTTTTTTGGGTGTACGAATGGCGCC	903
LOCUS		1812 bp DNA linear GSS 21-SEP-2004		Db	654	CACGGCGAGGATGTGAGTATGAATGGCAAGCTTTTCTGGGTTTACGTTGTGCTCC	713
DEFINITION		Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.		QY	904	ATTGACGTCGCTATATTGGGAAGTCTCTTGCTTTATTTCACTCAGCGCGAGGACGCG	963
ACCESSION	CL969864	GI:52394347		Db	714	CTTGGCATGTGTTATCTCTACCCCTGCTTTGTTTCTTCAAGCTCAGAAACATGG	773
VERSION	CL969864.1			QY	964	TGTTCAAGTGATAGGAGAACTCAAGAAAGGTTTGAATCCACCATCACTCACAATCTGCT	1023
KEYWORDS	GSS.			Db	774	CATTAGCATAATTGGGCAGCTCAATGTGGCCTGAATCGCCCTTCATGGGATAAATACT	833
SOURCE	Oryza sativa (indica cultivar-group)			QY	1024	ATTGTGTCGCTTACATGACTACAGCTGTCAAAACTGGCATGTGCTGGCATCATATC	1083
ORGANISM	Oryza sativa (indica cultivar-group)			Db	834	CTTTGATCCTCAATATTGGGCTCTCACTGTAAAGACTGGCCTTGTACCGGAATCATCTC	893
REFERENCE				QY	1084	ACTTCGGGAGCAATAGCAGTAGGAGAGCTTTGCCAATGTATATAAAATTTACAATATTC	1143
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.			Db	894	CCTAACCGAAGGAGTGGGAGATTTGCTTCACTCAAGGACTCACAGGATAGA	953
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis			QY	1144	TGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTCACTCTTG	1203
JOURNAL	Unpublished (2004)			Db	954	TGGTAACAAAGAGATGATGGCCATAGGATTGATGAACATTTGGGCTCATGTACATC	1013
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.			QY	1204	CTACCTCACAAAGAGACATTTTCGGTTCGGCTGTGAACATATAACGCTGGATGCAAGAC	1263
FEATURES				Db	1014	CTATGTAACTACAGGAGCAATCTCTCGTTCTCTGTGAACCAATGCTGTTGCAAGAC	1073
source	Location/Qualifiers			QY	1264	AGCAGCTTCCAACTTATTAATGTCACCTTGCAGTAAATGTTGACATTTGATTTCTCTGACAC	1323
	1. .1812			Db	1074	TGCCATGTCAAATGTGATCATGGCACTGACTGTCAATGTTGCTTCTCATGCT	1133
	/organism="Oryza sativa (indica cultivar-group)"			QY	1324	CTTGTTCATTAACACTCCCTCGTGTGCTATCAGCTATTTATCGTATCTGCAATGCTGG	1383
	/mol_type="genomic DNA"			Db	1134	ACTGTTTGTGTACACACCGAATGTTGTTCTTGGAGCAATCAATAATGCTGCTGATTGG	1193
	/db_xref="taxon:39946"			QY	1384	ACTCATAGATTAAGACAGCCATCCATCTATTTAAGGTTTGAACAAATTTGACTTTGCTGTG	1443
	/clone_lib="Oryza sativa Express Library"			Db	1194	GCTCATTTGATCTCCCTGCTGTATACAACTCTGGAAGATGGAACAAGATGATTTCTCTGT	1253
	/note="Oryza sativa exon trapped genomic sequences"			QY	1444	GTGCAATGATGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGC	1503
Query Match	23.3%; Score 570.8; DB 9; Length 1812;			Db	1254	GTGCTATGTGATTTGCTGTTGTTAATCTTATCTCAGTCCAAAGGCTTGCAATAGC	1313
Best Local Similarity	61.0%; Pred. No. 8.5e-138;			QY	1504	TATTTGTAATATCTGACTTTCGGGTACTTCTATTTTATTTGCAAGCCCAAGGACATTCGTTTT	1563
Matches	945; Conservative 0; Mismatches 597; Indels 6; Gaps 1;			Db	1314	GGTTGGTATATCAATATTTCAGGGTACTGTTGCAATCAACAGGCCAAAGATGATATCA	1373
364	GGGCATCAGTTATGCCAAGCTGCCAAGCTCCCTCCCAATCTTGGACTATATTCGAGCTT	423		QY	1564	GGGCAACATTCOAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACA	1623
174	GGGCATTAGCTACGCAAGCTGGCGAGCTTGCCCTCCGATCATCGGCTCTATTCGAGCTT	233		Db	1374	AGGGAACATCAAGGCACTGATTTTACCGAAATCTTATCATAGTACAAAGACGCTCAAG	1433
424	TATACCACCATTTGATTTATGCGATGATGGTAGCTCGAGGATTTTGGCAGTGGGACTGT	483		QY	1624	TGTTCTCGAATGCTAAATTTCTAGAGATTGATGCACCAATTTTACTTTGCCAATGCCAGCTA	1683
234	TGTGCCCGCATGTTGTACCGGTGCTGGGAGCTCGAGGACCTGCGCGTGGGCGGT	293		Db	1434	AGTTCTGCGTTCTTAATCTGACAGTGTAGGCTCTTATAAATCTTTGCAACACCAACTA	1493
484	GGCGTTTGGATCGCTTCTGATGGGTTGATGTTGAGTAATAGCGTTGATCCCAATGAAGA	543		QY	1684	TTTAAGAGAAGGATCACAGGTGGATTGATGAAGAAGAAGAAATTTAAAGCTACAGG	1743
294	GTGATCGGCTGCTCATTAATGGGTCATGCTGCGCAGCGGCTGAGCCCGCCGCGCA	353		Db	1494	CCTAAATGAAGGATTTAAAGATGGAT-----AGAAGAAGAAAGTTCTGCGAGGAACCA	1547
544	CCCAAGCTTTACCTCCAGCTGCTTTCACAGCTACATTTATGCTGTTGTTTTCAGGC	603		QY	1744	GGAGACTAGTTTGCAGTATGTTATTAATGATAGTCTGTTGGAAAACATTTGATACAG	1803
354	GCGCTGCTTCTCCTCCAGCTCGCCTTCACCTCCACTTCTTCCGCGGCTCGTCCAGGC	413		Db	1548	ACAAAAGTGAATCCCAATTTTCGTAATTTTGGATCTGTGAGTCTGCTGCAATTTGATACA	1607
604	TGCTTTGGGTCTGTTTATGTTGGGTTGATCGTGGATTTTCTGTACATGCAACCAATA	663		QY	1804	TGGAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTCT	1863
414	CTCCCTCGGCATCTCCGCTCGGCTTCATCATCGATTTCTGTCGAAGCGCGCTGCT	473		Db	1608	TGGCATATGTTTCTTCATCGACTTGAAGAAATCAACAGAGAAGCATGGCTAGAGCTTAT	1667
664	AGGTTTCATGGAGAGAGACCGACGCTGTTGCTGTGCAAGCACTAAAATCGATTTTGG	723		QY	1864	TTTGTGTCATCTGTAAAGTGAAGTGAAGAACTGAAGAACTGAACAAATCGAA	1911
474	CGGTTTCATGGCGCGCGCGCATCATCGTGTGCTGCGTCAGCAGCTCAAGGCGCTGCTCG	533		Db			
724	CCTTGAGCAATTTACCCATGGAGCTGATATCATATCAGTGATGCGCTCTGTTTTCACCA	783		QY			
534	CATCGTCCACTTACCACCCAGATGGGCTCGTCCCGCTCATGGCTCCGTCATCCACCA	593					

[illegible]

library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the clone ID of the original cDNA library that is also listed under 'OTHER EST'."

```
ORIGIN
Query Match      21.1%; Score 516.4; DB 7; Length 594;
Best Local Similarity 95.3%; Pred. No. 1.1e-123;
Matches 543; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 4 CGAGCTAGCTCGCACATTAAGTTATATAACACATA-TTTCGTTGCTTTAGAAATACTATTA 62
DB 570 CTAGCTAGCTCGCACATTTAATATATATACATATTTTCTAGCCAGCAAGAAATACTATCA 511
QY 63 TTGAAGATATGGGAGGTGATATAGTAGTACCTTTGGGCATGAACAACTTTGAGAGAG 122
DB 510 TTGAATATATGGGAGGTGATATAGTAGTACCTTTAGGCATGAACAACTTTGAGAGAG 451
QY 123 TGCACCAAGTCGAGGTTCACGCCACAGCGGTTTTCAGTCTCTAAAGTACTCTTTGA 182
DB 450 TGCACCAAGTGGAGGTTCACACCCTCAACCGTTTTTCAAGTCTCTTAAGTACTCTTTGA 391
QY 183 AGGAGACTTTCTTCCCTGTAGTACCTTTGAGGAGTTTCAAGAACAAAGCCAGCTTCCAAGA 242
DB 390 AGGAGACTTTCTTCCAGATGACCTTTGAGGAGTTTCAAGAACAAAGCCAGCTTCCAAGA 331
QY 243 AGTTCAATGCTGGCCTTCAGTCTTCTTCCCTCAATTTTCGAATGGGCTTCCAAATACACCT 302
DB 330 AGTTCAATGCTGGCCTTCAGTCTTCTTCCCTCAATTTTCGAATGGGCTTCCAAATACACCT 271
QY 303 TTCAGTTCTTGAAGCTGACTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTC 362
DB 270 TTCAGTTCTTCAAGCTGACTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTC 211
QY 363 AGGCGATCAGTTATGCCAAGCTCGCAACCTTCCCTCAATTTTCGAATGGGCTTCCAAATACACCT 422
DB 210 AGGCGATCAGTTATGCCAAGCTCGCAACCTTCCCTCAATTTTCGAATGGGCTTCCAAATACACCT 151
QY 423 TTATACCACCATTTGATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACTG 91
DB 150 TTATACCACCATTTGATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACTG 91
QY 483 TGGCGGTGATCGCTCTCATGGTTCGATGTTGAGTAATGCGTGTGATCCCAATGAAG 542
DB 90 TGGCAGTTGGATCGCTCTCATGGTTCGATGTTGAGTAATGCGTGTGATCCCAATGAAG 31
QY 543 ACCCAAGCTTTTACCTCCACCTGGCTTTCA 572
DB 30 ACCCAAGCTTTTACCTCCACCTGGCTTTCA 1
```

```
RESULT 11
CNS09YVW 2221 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTPEH70Z01 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX832405
VERSION BX832405.1 GI:42455111
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2221)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
```

```
REFERENCE 2 (bases 1 to 2221)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
FEATURES
source
Location/Qualifiers
1..2221
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/tissue_type="Hormone Treated Callus"
/plasmid="pQWVSPORT 6"
complement(1..2221)
/gene="At5g19600"
gene
ORIGIN
Query Match 20.9%; Score 512.6; DB 3; Length 2221;
Best Local Similarity 57.0%; Pred. No. 1.6e-122;
Matches 1050; Conservative 0; Mismatches 759; Indels 32; Gaps 5;

QY 168 TAAAGTACTCTTTGAGGAGACTTTCTTCCTGATGACCTTTTGAGGAGCTTCAAGACA 227
DB 164 TTAAGTCAAAATGTAAAGAAACCTTCTTCCCGATGATCCATTCAAACCGATATCGCAG 223
QY 228 AGCCA---GCTTCCAAAGAGTTTCATGCTTGGCCTTCAGTCTTCTTCTTCCCATTTTCCAAT 284
DB 224 AGCCAACCGTTTGCTAAACAAAGAGTATTGAGTATTTTTGCCCAATCTTTGAGT 283
QY 285 GGCTCCCAATACACCTTTTCAAGTCTTGAAGCTGACCTCATAGTGGCATCACCATCG 344
DB 284 GGCTCCCAAGTATGATATGCAAAATAAGTATGCTGCTCGCGGGAATCACCATCA 343
QY 345 CTAGCTTGGCCATCTCTCAGGCGATCAGTTATGCCAAGCTCGCAACCTCCCTCCATTC 404
DB 344 CCAGTCTTCTGTTCTCAGGCGATCAGCTATGCTAAGCTCGCCAGTATCCCTCCCATCA 403
QY 405 TTGGACTATATTCGAGCTTTATACCAACCATTTGATTTATGCGATGATGGGTAGCTCGAGGG 464
DB 404 TAGGCTTATATCAAGTTTGTGCGCCCTTTTGTGTATGCCGTGTTGCGAAGTTCCGACA 463
QY 465 ATTGGCAGTGGGAGCTGTGGCGTTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATG 524
DB 464 ATCTGCGAGTGGGAAACAGTCGCGAGCATGCTCTTTGCTGATCGCTGAGACATTCGGTGAGG 523
QY 525 CGGTTGATCCCAATGAAGACCCAAAGCTTTACCTCCACCTGGCTTTTACAGCTACATTAT 584
DB 524 AGAT---GATAAAAAATGAGCCTGAACTCTACCTTCATTTGATCTTCCACCGCCACTCTCA 580
QY 585 TTGCTCGTGTTTTTCAGGCTGCTTGTAGTTTGGGTTTGGTGTGATCGTGGATTTTC 644
DB 581 TCACCGGTCTATTCCAAATTTGCTATGGGATTTTGGGTTTGGGATATTTGGTGGATTTCT 640
QY 645 TGTTCATCAACCAATATAGGTTTCATGGGAGGAGCAGCCAGCGTGGTGTGTCTCGAGC 704
DB 641 TGTTCACACGACCAATAACGGGCTTCATGGGCGGTACCGCAATCATCTTCTCTCAAC 700
QY 705 AACTAAAATCGATTCTTTGGCTTTGAGCTTTACCCATGGAGCTGATATCATATCATGTA 764
```

Db 701 AGCTTAAGGCGCATCTTTGGACTAGTCCATTTTAAACGACAAAACCGATGTGATCAGTTC 760
QY 765 TGGCTCTGTTTTACCCAACTCATGAGTGGAGGTGGGAAGCTGTGTTAGGATGG 824
Db 761 TCACATCAATCTTAGACAACCGGGCTGAGTGGAAAGTGGCAAGTACCTTGGCTGGTGTCT 820
QY 825 TCTTCATTTTTCTTCCCTCTTAGCACAGATAC-TTCAGCAAAAACGACCAAGCTTTTTT 883
Db 821 GCITTTCTGTATTTTACAACTCCATCGATACATAAAGCAAGATACCCAAAGCTATTTT 880
QY 884 TGGGTGCAGCAATGGCGCAATTTGACGTCCGTTATATTTGGGAAGTCTCTTGGTTATTTTC 943
Db 881 GGGGTATCAGCAATGGGTCCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 940
QY 944 ACTCAGCCGAGAGACGCGTGTGAAGTGTAGGAGAACTGAAGAAGGCTTTGAATCCA 1003
Db 941 GTGAAGGAACAGACACAGGCATCTCTAGTTGGGCTTTAAAGAAAGACTAAATCCT 1000
QY 1004 CCATCACTCAAAATCTGGTATTTGTGCGCTTACATGACTACAGCTGTCAAAACTGGC 1063
Db 1001 CCTTCAT-TCACCTTTGAACCTTTGACTCCAAGTATCTAGGATGGTGTAAAGCGCGGA 1059
QY 1064 ATTGCTGGCATATACATCTTGGGAAGAAATAGCAGTAGGAAGAAGCTTTTGAATG 1123
Db 1060 ATCGTCACAGGACTCATCGCTTTGGCGGAAGGATAGCAATTTGGAAGAAGCTTTGCTGTA 1119
QY 1124 TATAAAATTTACAAATTTGATGCAACAAGAGATGATAGCTATTGGGACCATGAACGTA 1183
Db 1120 ATGAAGAAGCAAAACGGACGGGAACAAGAGATGATAGCTATTGGTCTTATGAACGTA 1179
QY 1184 GTTGGTTCTTTACCTCTTGTACTCTACACAGGACCAATTTTCGGGTTCGGCTGTGAAC 1243
Db 1180 ATTGGCTCCTTCACTTCTGCTATTGACAAACGGGCGCAATTTTCARAGACAGCAGTGAAC 1239
QY 1244 TATAACGCTGGATGCAAGACAGAGCTTCCAAATTTATATGTCATCTTGCAGTAATGTTG 1303
Db 1240 TACAACGGGGGCAAAAGACACCGATGTGCAATGTAGTAATGGGCGTTTGGCATGATGCTT 1299
QY 1304 ACATTTGTTATCTGTGACACCTTGTTCATTACACTCCCTGCTGTGCTATCAGCTATT 1363
Db 1300 GTGCTTCTCTTCTAGGCTCTATTGACGTACACACCATAGTGGTCTCTCAGGCCATC 1359
QY 1364 ATCGTATCTGCAATGCTTGGACTCATAGATTTATGAAGACCCATCCATCTATTTAAGGTT 1423
Db 1360 ATTATGTCAGCCATGTTGGGTCTCATCACTACAGAGAGATGTACCATCTCTTCAAGTT 1419
QY 1424 GACAAATTTGACTTTGTTGGTGCATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGT 1483
Db 1420 GACAAGTTTGATTTCTCGTCTGCATGTCGCGCTTTTTCGGTGTTCATTCTCTTAGCATG 1479
QY 1484 GAAATTTGGCTTAGTGCATAGCTATTGTAATATCTGTACTTTCGGTACTCTTATTTATGCA 1543
Db 1480 GACTACGGTCTCATCTCCGTTGGATTTTCTATCGTAAGACTTTGTTGACGTGTGA 1539
QY 1544 AGCCCAAGGACATTCGTTTTGGGCAACATTTCCAAATTTCTGTGATATACCGAAATTTGAG 1603
Db 1540 AGSCCTCCACTTGTAAATTTGGGAAGATACCAAACTCGGTTATTTTAGAGACATAGAA 1599
QY 1604 CACTATCAAAATGCAAAACATGTTCTCGAATGCTAAATTTCTAGAGATTTGATGACCAAT 1663
Db 1600 CAGTACCTGCTCCGAGGAGATGCTTGGTTATATCAATTTCTCAATAGGCTCTCCAGTC 1659
QY 1664 TACTTTGCCAATGCCAGCTATTAAAGAGAAAGGATCACAAGGTGGATTTGATGAAGAAGAA 1723
Db 1660 TTTTTTGGCAATAGCACTTACGTCGCGTGAACCGATCTTGAAGTGGATTCGTGATGAACCT 1719
QY 1724 GAAAGAAATTAAGCTACAGGGGAGACTAGTTTTCAGTATGTTATATAATTTGATATGAGTGT 1783
Db 1720 GAAGCCATCG-----AGTTTCTTCTCTCTGATCTCTCTGGA 1755
QY 1784 GTTGGAAACATTGATACAAGTGAATAGTATGCTTTGAAGAGGTGAAGAAGATTACAGAG 1843

Db 1756 GTTTCAACCATTTGACATGACCGGATGGAACACACTACTCGAAATTCAGAGAAATCTCTTGA 1815
QY 1844 AGAAGAGACTACAGCTTCTTTTGGTCAATCTGTAAAGTGAAGTGAATGAAGAACTGAAC 1903
Db 1816 TCAAAAACATCAAGATGGTGATATAAATCAAGATTTGAAGTCTTGGAAAGATGATG 1875
QY 1904 AAATCGAATTCCTCAAAATCAITTTAGGGAAGAAATGGATCTATCTGACTTTTGAAGAGGCC 1963
Db 1876 TTGTCCCAATTCGTGGAGAAGATAGGAAGAGTATATGTTCTTATCGATTGACGATGCA 1935
QY 1964 GTTGGAGCATCAACTTCAATCTACGTGCAAGCAAAAACGAA 2004
Db 1936 GTTCAAGCATGCCGATTTAATCTTACCACCACCAAGCCGGA 1976

CV130479 828 bp mRNA linear EST 03-SEP-2004
X9P02a03 Populus stem seasonal library Populus deltoides cDNA, mRNA
sequence.
CV130479
CV130479.1 GI:51874399
EST.
Populus deltoides
Populus deltoides
ORGANISM
SOURCE
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Park, S. and Han, K.-H.
Gene expression profile during seasonal growth cycle in poplar tree
Unpublished (2003)
Contact: Kyung-Hwan Han
Department of Forestry
Michigan State University
126 Natural Resources, East Lansing, MI 48824-1222, USA
Tel: 517 353 4751
Fax: 517 432 1143
Email: hanky@msu.edu.

FEATURES
source
1..828
/organism="Populus deltoides"
/mol_type="mRNA"
/strain="ILL-129"
/db_xref="taxon:3696"
/tissue_type="stem"
/dev_stage="1 year old"
/clone_lib="Populus stem seasonal library"

ORIGIN
Query Match 20.5%; Score 501.4; DB 7; Length 828;
Best Local Similarity 77.1%; Pred. No. 1e-119;
Matches 610; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 1078 CATATCATTGCGGAAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATATAAAATTTACAA 1137
Db 37 CATTCGGCGGGGAAGGAATAGCAGTAGGAAGAAGTTTGGCATGTTCAAGAAATTTACCA 96
QY 1138 TATTGATGGCAACAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTCTTAC 1197
Db 97 TATAGATGGCAACAAGAGATGATGCAATTTGGGACCATGAACATTTGAGGCTCTTGCAC 156
QY 1198 CTCTTGCTACCTCACAAACAGGACCAATTTTCGGGTTCGGGTGTAACATAAAGCTGGATG 1257
Db 157 CTCTTGCTATCTCAACAACAGGGCCATTTCCGGATCGGCTGTGAATTTCAATTCAGGATG 216
QY 1258 CAAGACAGCAGCTTCCAAATTTAATGTCACTTCAGTGAATGTGTAATTTGATATTCCT 1317
Db 217 CAAGACAGCAGTATCGAATTTGTATGTCGCTTGGCAGTGTGGTCACTGTTGTTATTCCT 276
QY 1318 GACACCTTGTTCATATACCTCCCTGCTGCTATCAGCTATATCGTATCTGCAAT 1377
Db 277 AAGCGCAATGTTCCCATTTACACTCCACTTGTGTTCTTATCTCTATTTATCTCTGCCAT 336

QY	1378	GCTTGGACTCATGATTATGAAGCAGCGCATCCATCTATTAAAGTTGACAAATTTGACTT	1437
Db	337	GCTAGGCTTTATAGATTATGAAGCAGCAATCCATCTTTGGACTGTCTGACAAAGTTGCACTT	396
QY	1438	TGTGGTGTGATGATGATCATACATTGGCGTGGTCTTTGGCAGTGTGTGAATTCGCTTAGT	1497
Db	397	CATCGTGTGATGAAGCCATATGCTGTGTGGTTTTTTGGCAGTGTGAGATTGGCTTAGT	456
QY	1498	CATAGCTATTGTAAATATCTGTACTTTCGGGTACTTCTATTATTATTCGAAGGCCAAGGACATT	1557
Db	457	CATCGCGTTGGCAATTTCTTTGCTCAGATTACTTCTCTCTGTCGCAAGACCAAAACATT	516
QY	1558	CGTTTGGGCAACATTTCCAAATCTGTGATATATACCGGAATGTTGAGCAGCTATAAATGC	1617
Db	517	TATTTCTTTGGAAACATTTCCAAATTCGATGATCTACAGAAATGTTTGAACAGTATACGAAC	576
QY	1618	AAAAACATGTTTCTTGGAAATGCTAATCTTAGAGATTGATGCACCAATTTTACTTTGCCAATGC	1677
Db	577	AAGCAGGTTCCCGGTGTTCTCATCTCAGAGATTGATGCTCTTACTTTGCAAAATGC	636
QY	1678	CAGCTATTTAAGAGAAAGGATCAACAGGTGGATTGATGAAGAAGAGAAAGAAATTAAGC	1737
Db	637	AAGCTACTTAAAGAGAAAGGATTGCGAGGTGGTTGATGAAGAGAGAGAAACAGTTAAATC	696
QY	1738	TACAGGGGACATGATTGTCAGTATGTTATAATTTGATATGAGTGTCTGTTTGGAAACATTGA	1797
Db	697	TTCAGGAGAAACCCAGCGCTGAGTATGTTATCTAGACATGGAGCGGTAGTTAATCTGA	756
QY	1798	TACAGTGGAAATAGTATGCTTCAAGAGGTGAGAGAGATTACAGAGAGAGAGAGTACA	1857
Db	757	CACAAGTGGGATTGGCATGCTTGGAGAAAGTAAGAAAGTAATGGACAGAGAGGAACTCAA	816
QY	1858	GCTTGTGTTTGG	1868
Db	817	GCTTGCTTGG	827
RESULT 13			
BI972185			
LOCUS			
DEFINITION			
eag88c07.y1 Gm-cl084 Glycine max cDNA clone GENOME SYSTEMS CLONE			
ID: Gm-cl084-1262.5' similar to TR:064435 O64435 SULFATE			
TRANSPORTER. [3] TR:065025 TR:023250 ;, mRNA sequence.			
BI972185			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Glycine max (soybean)			
Glycine max			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;			
Glycine.			
1 (bases 1 to 498)			
Shoemaker, R., Kaim, P., Vodkin, L., Erpelding, J., Coryell, V.,			
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,			
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,			
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,			
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,			
McCann, R., Waterston, R. and Wilson, R.			
Public Soybean EST Project			
Unpublished (1999)			
Contact: Shoemaker R/Public Soybean EST Project			
Public Soybean EST Project			
Washington University School of Medicine			
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA			
Tel: 314 286 1800			
Fax: 314 286 1810			
Email: est@watson.wustl.edu			
When it has been determined, an EST from the other end of this			
clone is listed in the 'Other ESTs on clone' field. This clone is			
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD			
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)			

High quality sequence stop: 421.			
Location/Qualifiers			
1. .498			
/organism="Glycine max"			
/mol_type="mRNA"			
/cultivar="Williams 82"			
/db_xref="taxon:3847"			
/clone="GENOME SYSTEMS CLONE ID: Gm-cl084-1262"			
/tissue_type="Etiolated hypocotyls (Williams 82)"			
/lab_host="DH10B"			
/clone_lib="Gm-cl084"			
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed by M. Bhattacharyya from mRNA isolated from etiolated hypocotyls from the cultivar Williams 82. Tissue was inoculated with Phytophthora sojae race 1 and tissues were harvested 2 and 4 hours following infection. The library is the pool of these two time points. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by M. Bhattacharyya in the laboratory of Dr. Randy Shoemaker at Iowa State University."			
ORIGIN			
Query Match 20.1%; Score 493.2; DB 4; Length 498;			
Best Local Similarity 99.4%; Pred. No. 1.2e-117;			
Matches 495; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	1656	CACCAATTTACTTTGGCAATGCCAGCTATTTAAGAGAAAGGATCAAGGTGGATTGATG	1715
Db	1	CACCAATTTACTTTGGCAATGCCAGCTCAGTAAGAGAAAGGATCAAGGTGGATTGATG	60
QY	1716	AAGAAGAGAAAGAAATTAAGCTACAGGGGAGACTAGTTTGCAGTATGTTATATTGATA	1775
Db	61	AAGAAGAGAAAGAAATTAAGCTACAGGGGAGACTAGTTTGCAGTATGTTATATTGATA	120
QY	1776	TGAGTGCTCTTGGAAACATTGATACAGTGGAAATAGTATGCTTGAAGAGGTGAAGAAGA	1835
Db	121	TGAGTGCTCTTGGAAACATTGATACAGTGGAAATAGTATGCTTGAAGAGGTGAAGAAGA	180
QY	1836	TTACAGAGAGAGAGAGCTACAGCTGTTTGGTCAATCCTGTAAGTGAAGTGAATGAAGA	1895
Db	181	TTACAGAGAGAGAGAGCTACAGCTGTTTGGTCAATCCTGTAAGTGAAGTGAATGAAGA	240
QY	1896	AAGTGAACAAATCGAAGTTCCAAAATCATTTAGGGAGAAATGGATCTATCTGACTGTTG	1955
Db	241	AAGTGAACAAATCGAAGTTCCAAAATCATTTAGGGAGAAATGGATCTATCTGACTGTTG	300
QY	1956	AAGAGCCGCTGGAGCATGCAACTCTCAATCTACGTGCAAGCAAAAGCAACCAAGAAAG	2015
Db	301	AAGAGCCGCTGGAGCATGCAACTCTCAATCTACGTGCAAGCAAAAGCAACCAAGAAAG	360
QY	2016	ATGAACACAGAGGGTTGGAACAAATGTGTGACTCAGTCAATATGCGCAAGAGATTTCTAAATA	2075
Db	361	ATGAACACAGAGGGTTGGAACAAATGTGTGACTCAGTCAATATGCGCAAGAGATTTCTAAATA	420
QY	2076	ACTCAAAAAGCTTATTCGTTTTTCGTTTCTTAGTAAATGTTTACCACCTACAAATGTGGGCATGAG	2135
Db	421	ACTCAAAAAGCTTATTCGTTTTTCGTTTCTTAGTAAATGTTTACCACCTACAAATGTGGGCATGAG	480
QY	2136	AATTTCTGAATCACGCCG	2153
Db	481	AATTTCTGAATCACGCCG	498
RESULT 14			
CK269206			
LOCUS			
CK269206			
982 bp			
mRNA			
linear			
EST 03-AUG-2004			

```

DEFINITION EST715284 potato abiotic stress cdna library Solanum tuberosum cDNA
clone POACL46 5' end, mRNA sequence.
ACCESSION CK269206
VERSION CK269206.1 GI:39826184
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 982)
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST715285
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source            1..982
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POACL46"
                     /tissue_types="abiotic stress treated leaf and root tissue"
                     /lab_host="DH10B-Tona"
                     /notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
                     supplier: Solanum tuberosum var. Kennebec plants were
                     grown from cuttings on a 16hr light/8 hr dark cycle at 25
                     C for 3-4 weeks. Abiotic stress conditions were applied to
                     four separate sets of plants. Set 1 involved saturation of
                     the soil with 150 mM NaCl and tissues were harvested at
                     following application of the salt stress (leaves: 2hr,
                     6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
                     Set 2 were grown under the standard conditions and then
                     were water stressed by withdrawal of further watering
                     applications. Drought stressed plants were harvested after
                     cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
                     and 5d). Set 3 were grown under the standard conditions
                     and then were cold stressed by placement at 4 C. Cold
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d. Set 4 were grown under the standard conditions and
                     then were heat stressed by placement at 35 C. Heat
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d and 4d and heat-stressed roots were harvested at 6 hr,
                     12 hr, 1 d, and 4d. RNA was isolated from all tissues and
                     equal RNA from each tissue and stress was pooled to
                     construct the cDNA library. RNA sample."
ORIGIN
Query Match          19.4%; Score 476; DB 7; Length 982;
Best Local Similarity 71.3%; Pred. No. 4.8e-113;
Matches 626; Conservative 0; Mismatches 250; Indels 0; Gaps 0;
QY 123 TGCACCAAGTCGAGGTTCCACCCGACAGCCGCTTTTCAAGTCTCTAAAGTACTCTTTCA 182
Db 101 TACACAGAGTTGAATTCCTCCACACAGCCCTTTTCAAGTCATTGAAAATACAGTGA 160
QY 183 AGAGACTTTCTTCCTGTATGACCTTTTGAGGAGTTCAGAACAGCAGCTTCCAGA 242
Db 161 AGGAACAATTAATTCAGATGATCCACTTAAGCAATTCAGAACCAAAAACCTTTGAGAA 220
QY 243 AGTTTCATGCTGGCCCTTCAGTCTCTTCCCAATTTTCGAATGGGCTCCCAATACCT 302
Db 221 AATTATACTTGGAGTTCAAATATTTTTCCAAATTTTGAATGGGCTTCAAGATACAAAT 280

```

```

QY 303 TTCAAGTTCTTTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTC 362
Db 281 TTGGTTTCTTTCAAGCGCAGATCTTATTTGCTGGAAATACATAGTAGTCTTGTCTATTCCTC 340
QY 363 AGGCAATCAGTTATGCAAGCTCGCCAACTCCCTCCCAATTTCTTGGACTATATTCAGCT 422
Db 341 AAGGAATAAGCTATGCTAAGCTGGGAACTTGCCACCTATCTACTTGGCTTATATTCAGTT 400
QY 423 TTATATCCACCATTTGATTTATGATGATGATGCTGAGGATTTGGCAGTGGGAGCTG 482
Db 401 TTATTTCCACCATTTAGTGATGATGATTAATGGAAGTTCAAGAGATTTAGCAGTGGGAACAG 460
QY 483 TGGCGGTTGGATCGCTTCTTGATGGGTTCCATGTTGAGTAATGCGCTTGATCCCAATGAAG 542
Db 461 TTGCTGTTGGATCTCTTCTCATGGCTTCTATGTTGGAGCTGAAGTTAACCTGCTGAAA 520
QY 543 ACCCAAGCTTTACCTCCACCTGGCTTTCACAGCTACATTTATTTGCTGTGTTTTCAGG 602
Db 521 ATCCAACACTTTATCTTCATCTTGTCTTTCACGCCACTTCTTTCACCGGTTTATTCGAAG 580
QY 603 CTGCGCTTGGGCTGTTTGTAGGTTGGGTTGATCGTGGATTTTCTGTACATGCAACCATAA 662
Db 581 TAGCACTCGGATTTTTCAGGCTGGGATTTATAGTGGATTTTCTATCACATGCAACAATTG 640
QY 663 TAGGGTTTCATGGGAGGAGCAGCCAGCTGGTGTGTGTCAGCAAACTAAATCGATCTTTG 722
Db 641 TGGGATTTATGGAGGAGCAGCTACAGTAGTGATGATCTCCAGCAATTAAGGAATATTG 700
QY 723 GCTTTGAGCAATTTACCCATGAGCTGATATCATATCATGATGATGCTGTGTTTTCACCC 782
Db 701 GTCTTGAAACATTTTACTCATGCACTGATGTTGTCTCTCTGTCTGTCTGTATTTTTCAC 760
QY 783 AAACATCATGAGTGGAGTGGGAAAGTGTGTGTAGGATGTCTTCAATTTTCTTCTCTCC 842
Db 761 AAACATCACCGTGGCGATGGGAAAGTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 820
QY 843 TTAGCACAAAGATACTTTCAGCAAAAACGACCAAGGTTTTTTTTGGGTGTGAGCAATGGCGC 902
Db 821 TGCTCGCCAAATTTTGTAGCCAGAGAGACCAAAAGCTGTCTGTGGTATCGGCAATGGCGC 880
QY 903 CATTGACGCTCGTTATATTTGGGAAGTCTCTTGGTTATTTTCACTACGCGGAGAGCAG 962
Db 881 CATTGACGCTCTCATATTTGGGAATCTATTTCTCGTTTATATAACGCATGCTGAAAAACAG 940
QY 963 GTGTTGAAGTGTAGGAACTGAAGAGGGTTTGA 998
Db 941 GGTGTTGCTGTGATTCGAGAGCTGAAGAAAGGATAA 976

RESULT 15
BU090662
LOCUS
DEFINITION
ACCESSION BU090662
VERSION BU090662.1 GI:22540819
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE 1 (bases 1 to 583)
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,
            Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
            Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
            Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
            Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
            McCann, R., Waterson, R. and Wilson, R.
            Public Soybean EST Project
            Unpublished (1999)

```

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
High quality sequence stop: 420.

FEATURES

Location/Qualifiers

1. .583
/organism="Glycine max"
/mol_type="mRNA"
/cultiivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl066-830"
/tissue_type="Leaf and shoot tip, salt stressed, 2 week old seedling"
/lab_host="DH10B"
/clone_lib="Gm-cl066"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from unexpanded leaves and the shoot tips of 2 week old seedlings from the cultivar Williams. The 2 week old seedlings were salt stressed in a solution of 500mM NaCl for 3 days prior to harvesting. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 19.1%; Score 468.2; DB 5; Length 583;
Best Local Similarity 90.3%; Pred. No. 4.5e-111;
Matches 51; Conservative 0; Mismatches 54; Indels 1; Gaps 1;

QY 4 CGAGCTAGCTCGCACATTAAGTTATATACACATA-TTCTGCTTGCTTAGAAATACTATTA 62
DB 18 CTAGTAGCTCGCACATTTAATTATATTACACATATTTTCTAGCCAAAGAAATACTATCA 77
QY 63 TTGAAGATATGGGAGGTAGATATAGTACCTTTGGGCATGAACAACCTTTGAGAGAG 122
DB 78 TTGAATATATGGGAGGTAGATATAGTACCTTTTAGGCATGAACAACGTTGAGAGAG 137
QY 123 TGCACCAAGTCGAGGTTCCACGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGA 182
DB 138 TGCACCAAGTGGAGGTTCCACACCTCAACCGTTTTTCAAGTCTCTTAAGTACTCTTTGA 197
QY 183 AGGAGACTTTCTTCCCTGATGACCTTTTGAGGAGTTTCAAGAACAGCCAGCTTCCAAGA 242
DB 198 AGGAGACTTTCTTCCAGATGACCTTTTGAGGAGTTTCAAGAACAGCCAGCTTCCAAGA 257
QY 243 AGTTTCATGCTGGGCTTCAGTTCTTCTCCCAATTTTCGAATGGGCTCCCAATACACCT 302
DB 258 AGTTTCATGCTGGGCTTTAGTACTTCTTCCCAATTTTCGAATGGGCTCCCAATACACCT 317
QY 303 TTCAAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCCTC 362
DB 318 TTCAAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCCTC 377
QY 363 AGGCAATCAGTTATGCAAGCTCGCCCAACCTCCCTCCAATTTCTTGGACTATATTCGAGCT 422
DB 378 AGGCAATCAGTTATGCAAGCTCGCCCAACCTCCCTCCAATTTCTTGGACTATATTCGAGCT 437
QY 423 TTATACCACCAATGATTTATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAGCTG 482

Db 438 TTATACCACCAATGATCTATGCGATGATGGGTAGCTCGAGGGATTTATGCGAGGGGACTG 497
QY 483 TGGCGGTTGGATCGCTTCTGATGGGTTCCGATGTTGAGTAATGCGCTTGTATCCCAATGAAG 542
DB 498 GGGCAGTTTGAAGGCTTCTTATGGGGTTCGATGTTGATGCGGTCGCTCCCTATGAAG 557
QY 543 ACCCAAAGCTTTACCTCCACCTGGCT 568
DB 558 ACGCCATGCTTTACCTCCACCTTCGCT 583

Search completed: August 31, 2005, 08:45:56
Job time : 7928 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2005, 03:36:12 ; Search time 122 Seconds

(without alignments)
2155.714 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGWNV 680

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3475	100.0	680	3 AAY44943	Aay44943 Soybean s
2	2641.5	76.0	658	3 AAG32200	Aag32200 Arabidops
3	2638.5	75.9	658	3 AAG31621	Aag31621 Arabidops
4	2638.5	75.9	703	3 AAG31620	Aag31620 Arabidops
5	2164	62.3	579	3 AAY44936	Aay44936 Corn sulp
6	2070	59.6	510	3 AAG32201	Aag32201 Arabidops
7	2067	59.5	510	3 AAG31622	Aag31622 Arabidops
8	1901	54.7	621	3 AAY44942	Aay44942 Soybean s
9	1898.5	54.6	631	3 AAG329247	Aag329247 Arabidops
10	1855	53.4	457	3 AAG32202	Aag32202 Arabidops
11	1843.5	53.1	688	3 AAY44935	Aay44935 Corn sulp
12	1807.5	52.0	656	3 ABP53777	Abp53777 Arabidops
13	1798.5	51.8	656	3 AAY44944	Aay44944 Wheat sul
14	1444.5	41.6	499	3 AAG329248	Aag329248 Arabidops
15	1271.5	36.6	446	3 AAG329249	Aag329249 Arabidops
16	961.5	27.7	593	3 AAY44945	Aay44945 Wheat sul
17	731.5	21.1	226	3 AAY44938	Aay44938 Artichoke
18	706.5	20.3	735	7 ADM56357	Adm56357 Frog SLC2
19	687.5	19.6	596	7 ABOJ5066	Abpj5066 Pseudomon
20	679.5	19.6	735	7 ADM56276	Adm56276 Mouse SLC
21	679.5	19.6	758	7 ADM56274	Adm56274 Mouse SLC
22	675.5	19.4	311	3 AAY44937	Aay44937 Corn sulp
23	670	19.3	753	7 ADM56359	Adm56359 Pig SLC26
24	668.5	19.2	758	7 ADM56270	Adm56270 Human SLC
25	668	19.2	738	7 ADM56272	Adm56272 Human SLC

26	665.5	19.2	758	3 AAY71067	Aay71067 Human mem
27	664.5	19.1	698	8 ABM84951	Abm84951 Human dia
28	663	19.1	697	8 ABM84952	Abm84952 Human dia
29	662	19.1	751	4 AAG67162	Aag67162 Amino aci
30	662	19.1	4115	5 ABP53570	Abp53570 Human NOV
31	658.5	18.9	746	3 AAY44358	Aay44358 P. chryso
32	657	18.9	721	6 ABG75831	Abg75831 Transpore
33	651.5	18.7	328	5 AAE22124	Aae22124 Human 543
34	651.5	18.7	328	6 ADA20211	Ada20211 Sulphate
35	650.5	18.7	744	7 ADM57787	Adm57787 Murine pr
36	650.5	18.7	744	8 AD86098	Ad86098 Murine pr
37	650	18.7	744	7 ADM57785	Adm57785 Gerbil pr
38	650	18.7	744	8 AD86096	Ad86096 Gerbil pr
39	649.5	18.7	716	8 ABM84960	Abm84960 Human dia
40	647.5	18.6	739	2 AAY08929	Aay08929 Rat PrDST
41	647.5	18.6	739	7 ADD46847	Add46847 Rat Prote
42	643.5	18.5	688	8 ABM84953	Abm84953 Human dia
43	642.5	18.5	739	5 AAU74765	Aau74765 Human SLC
44	641.5	18.5	739	2 AAY08928	Aay08928 Human DTD
45	641.5	18.5	739	5 AAU74762	Aau74762 Human SLC

ALIGNMENTS

RESULT 1

AAAY44943

ID AAY44943 standard; protein; 680 AA.

XX AAY44943;

XX 23-MAY-2000 (first entry)

XX Soybean sulphate permease-2.

XX Sulphate Permease; sulphate assimilation protein; soybean; probe;
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
KW antibody; screen.

XX Glycine max.

XX WO200004154-A2.

XX 27-JAN-2000.

XX 13-JUL-1999; 99WO-US015810.

XX 14-JUL-1998; 98US-0092833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Allen SM, Falco SC, Thorpe CJ;

XX WPI; 2000-195025/17.

XX N-PSDB; AAY50489.

XX Nucleic acid fragments encoding sulfate assimilation proteins in plants
PT and seeds useful as probes for isolating cDNAs and genes encoding
PT homologous proteins, in producing transgenic plants.

XX Claim 6; Page 53-55; 79pp; English.

XX The present amino acid sequence is the soybean sulphate permease, a
CC sulphate assimilation protein. This sequence is obtained from sf11
CC library, clone sf11.pk0043.g10, derived from soybean immature flower. It
CC has 75% sequence identity to Arabidopsis thaliana sulphate permease (gi
CC 2285885). This sequence is used as a probe to isolate other plant
CC sulphate assimilation proteins, for genetic and physical mapping of
CC related genes and as markers of traits linked to the gene. This is useful
CC for plant breeding and to construct chimeric genes, used to create
CC transgenic plants with altered levels of sulphate permease. The sulphate
CC permease peptides are useful for producing antibodies, that are used to
CC screen and isolate cDNA clones

PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145114P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 08-OCT-1999; 99US-0157865P.
PR 08-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 76.0%; Score 2641.5; DB 3; Length 658;
Best Local Similarity 75.9%; Pred. No. 2e-242;
Matches 500; Conservative 81; Mismatches 75; Indels 3; Gaps 3;

Qy 24 MGSVDYEPYPLGMNFRVHQ-VEVPPQPPFFKSLKYSLKETPPDDPDLROFKKPKASKKF 82
Db 1 MGTEDYTFQCAEELHRRHHTVEAPQPFLKSLQYSVKETLPPDDPDPFROFKQNASKF 60
Qy 83 MLGLQFFFPFIFEWAPKTYTFQFLKADLIAGTTIASLAIPOGISYAKLANLPPIILGLYSSEI 142
Db 61 VLGLKYFLPIFEWAPRYNLKFFKSDLIAGTTIASLAIPOGISYAKLANLPPIILGLYSFV 120
Qy 143 PPLIYAMGSSRDLAGVTAVAGSLLMGSMLSNAVDNPDPKLYLHLAFTATLTFAGVFQAA 202
Db 121 PPLVAVLGSSRDLAGVTAVASLLTGAMLSKEVDAEKOPKLYLHLAFTATTFAGVLEAS 180
Qy 203 LGLFRLGLIVDFLSHATIIIGFMGAATVVCLOQLKSLILGLEHETHGADIISVNRSVPTQT 262
Db 181 LGIFRLGFIIVDFLSHATIVGFMGAATVVSLOQLKGIFGLKHFSTDVSVNRSVFSQT 240
Qy 263 HEWRWESAVLGCVFIIFLLSTRYFSKKRPRFFWVSAMAPLTSVILGSLAVYFTHAEKHGV 322
Db 241 HEWRWESVGLGCGFLFLLSTRYFSIKKPKFFWVAAMAPLTSVILGSLAVYFTHAEKHGV 300
Qy 323 EVIGELKKGLNPPSLTNLVFVSPYMTTAVKTVGIWVGIISLAEGIAVGRSPAMTKYNNIDG 382
Db 301 QVIGDLKKGLNPLSGSDDLFTSPYMTAVKTVGIITGIIAALAEVAVGRSPAMTKYNNIDG 360
Qy 383 NKEMIAIGTNVNVGTSFTCYLTTGPFPSAVNNACKTAASNIINSLAVMLLFLFTPL 442
Db 361 NKEMIAFGMMNIVGTSFTCYLTTGPFPSAVNNACKTAASNIINSLAVMLLFLFTPL 420
Qy 443 FHYTPLVLSAIIIVSAMLGLIDYEAAILHFLKVDKDFVVCMSAYIGVVFSGSGLVIAI 502

Db 421 FHYTPLVLSAIIISAMLGIDYQAAIHLWKVDKDFLVCMSAYGVVFGSVEIGLWAV 480
Qy 503 VISLRLVLLFIAPRPTFVLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYL 562
Db 481 AISIARLLLVSRPKTAVKGNIPNSMIYRNTQYPPSRTVPGILILEIDAPIYFANASYL 540
Qy 563 RERITRWIDEEERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVL 622
Db 541 RERIIRWIDEEERVKQSGESSIQYIILDSAVGNIDTSGISMMVLEIKKVIDDRALKVL 600
Qy 623 VNPVSEVMKLNKSKF-QNHLGKKWYLTVEEAVGACNFNLASKNPNPKDTEGWNV 680
Db 601 SNPKGEVVKLTRSKFTGDHKGWMLFTVGEAVEACSYMLHTFKTEP-ASKNEPNWV 658

RESULT 3
AAG31621
ID AAG31621 standard; protein; 658 AA.
XX AC AAG31621;
XX DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38004.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 03-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.

101-JUN-1999; 99US-0137222P.
03-JUN-1999; 99US-0137528P.
04-JUN-1999; 99US-0137502P.
07-JUN-1999; 99US-0137724P.
08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139899P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
29-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
01-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.
19-JUL-1999; 99US-0144332P.
19-JUL-1999; 99US-0144333P.
19-JUL-1999; 99US-0144334P.
19-JUL-1999; 99US-0144335P.
20-JUL-1999; 99US-0144352P.
20-JUL-1999; 99US-0144632P.
21-JUL-1999; 99US-0144884P.
21-JUL-1999; 99US-0144814P.
21-JUL-1999; 99US-0145086P.
21-JUL-1999; 99US-0145088P.
22-JUL-1999; 99US-0145085P.
22-JUL-1999; 99US-0145087P.
22-JUL-1999; 99US-0145089P.
22-JUL-1999; 99US-0145192P.
23-JUL-1999; 99US-0145145P.
23-JUL-1999; 99US-0145218P.
23-JUL-1999; 99US-0145224P.
26-JUL-1999; 99US-0145276P.
27-JUL-1999; 99US-0145913P.
27-JUL-1999; 99US-0145918P.
27-JUL-1999; 99US-0145919P.
28-JUL-1999; 99US-0145951P.
02-AUG-1999; 99US-0146386P.
02-AUG-1999; 99US-0146388P.
02-AUG-1999; 99US-0146389P.
03-AUG-1999; 99US-0147038P.
04-AUG-1999; 99US-0147204P.
04-AUG-1999; 99US-0147302P.

PR	05-AUG-1999;	99US-0147182P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148694P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0151438P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 75.9%; Score 2638.5; DB 3; Length 658;		
Best Local Similarity 75.7%; Pred. No. 3.8e-242;		
Matches 499; Conservative 82; Mismatches 75; Indels 3; Gaps 3;		
Qy	24	MGSDVDEYPLGMNFRVHQ-VEVPPPPQPFKSLKYSLKETPPDDPLRQFKNKPASKKF 82
Db	1	MGTEDYTFQGAELHRRHHTVEAPQPFKLKSLQYSVETLPPDDPFRQFKQNSRKF 60
Qy	83	MLGLQFFFFIFEWAPKYTFQFLKADLIAGITIASLAIPQGISYAKLANLPPIILGLYSSFI 142
Db	61	VLGLKYFLPIFEWAPRYNLKFFKSDLIAGITIASLAIPQGISYAKLANLPPIILGLYSSFV 120
Qy	143	PPLIYAMGSSRDIAVCTVAVGSLLMGSMLSNADNEDPKLYLHLAFTATLACGVFOAA 202
Db	121	PPLVYAVLGGSSRDIAVGTVAASLLTGAMLSKEVDKOPKLYLHLAFTATTFAGVLEAS 180
Qy	203	LGLFRLGLIVDFLSHATIIIGFMGGAATVVCLOQKSLILGLEHFTHGADIIISVMSRVSFTQT 262
Db	181	LGIFRLGFIIVDFLSHATIIIGFMGGAATLVSLQKLGIFGLKHFTDSTDVISVMSRVSFTQT 240
Qy	263	HEWRWESAVLGCVFIFFLLLSTRYFSKKRPRFPFVWSAMAPLTSVILGSLLVYFTHAEKHGV 322
Db	241	HEWRWESVGLGCGFLFLLSTRYFSIKKPKFFWVAAMAPLTSVILGSLLVYFTHAEKHGV 300
Qy	323	EVIGELKGLNPPSLTNLVFVSPYMTAVTKTAVTGIVGIIISLAEGIAVGRSPAMYKNYNIDG 382
Db	301	QVIGDLKGLNPLSGSLDIFTSPYMTAVTKTGLITGIIALAEVAVGRSPAMFKNYNIDG 360
Qy	383	NKEMIAIGTMNVVGSFTSCVLTTPFSPSAVYNAGCKTAASNIIMSLAVMLTLLFLTPL 442
Db	361	NKEMIAFGMMNIVGSFTSCVLTTPFSPSAVYNAGCKTAMSNIVMAIWMFTLLFLTPL 420
Qy	443	FHYTPLVLSAIIVSAMGLIDYEAHILFKVDKDFVVCMSAYIGVFGSVEIGLVIAI 502
Db	421	FHYTPLVLSAIIISAMGLIDYQAAHILKWKDKDFLVCMSAYVGVFGSVEIGLVAV 480
Qy	503	VISVLRVILFIARPTFVLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYL 562
Db	481	AISIARLLLLFVSRPKTAVKNIPNSMIYRNTQYPSRTPVPGILILEIDAPIYFANASYL 540
Qy	563	REIRTWIDEEERIKAETSLQYVVIDMSAVGNIDTSGISMLEEVKKITERRELQVL 622
Db	541	REIRTWIDEEERVQSGESSLOYIILDMASVGNIDTSGISMMVEIKKVIDRALKVL 600
Qy	623	VNPVSEVMKKLANKSF-QNHLGKKIYLYTVEEAVGACNPNLRASKTNPKKDETEGWNV 680
Db	601	SNPKGEVWKLTJRSKFIDGHLGKEMFLITVGEAVEACSYMLHTFKTEP-ASKNEPWNV 658
RESULT 4		
AAG31620		
ID	AAG31620 standard; protein; 703 AA.	
XX	AC AAG31620;	
XX	AC AAG31620;	
DT	17-OCT-2000 (first entry)	
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 38003.	
XX	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX	Arabidopsis thaliana.	
XX	EP1033405-A2.	
PN	06-SEP-2000.	
PD	25-FEB-2000; 2000EP-00301439.	
PF	25-FEB-1999; 99US-0121825P.	
XX	05-MAR-1999; 99US-0123180P.	

PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0133425P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-014342P.
PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145911P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147933P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148344P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.

Qy	225	GGAAVVCLQQLKSIIGLEHFTGADIISVMRSVFTQTHEWRWESAVLGCVFIFLLSTR	284	99US-0132485P.
Db	123	AGAAVVCLQQLKGLVHFTTSTDVVSVMESVFQTHQWRWESVLGCGPLFLLVTR	182	99US-0132486P.
Qy	285	YFSKKRPRPFWSAMAPLTSVIIIGSLVYTHAEKGVVEVIGELKGLNPPSLTNLVFS	344	99US-0132487P.
Db	183	FISKRRPKLFWISAAAPLTSVWILGSLVLYLTHAENHGEIVGYLKKGLNPPSVTSLOFSP	242	99US-0132863P.
Qy	345	PYMTAVKTIQVVGIIISLAGIAGVGSFAMYNKYNIDGNKEMIAIGTMNVVGSFTSCYLT	404	99US-0134256P.
Db	243	PYMWLAKTGIITGVIALAGIAGVGSFAMFKYHMTDNKEMIAIGTMNVLGLTSCYLT	302	99US-0134258P.
Qy	405	TGPFSSAVNNAGCKTAASNIIMSLAVMLTLFLTPHLYTFLVVLISAIIVSAMLGLID	464	99US-0134218P.
Db	303	TGPFSSAVNNAGCTAMSNVMSLAVMVTLLFLTPHLYTFLVVLISAIIVSAMLGLVD	362	99US-0134219P.
Qy	465	YEAAILHFKVDKDFVVCMSAYIGVVGSGVEIGLVAIVISLVRVLLFIARPTTFVLGNI	524	99US-0134221P.
Db	363	FGAALHLWRVDKDFVCAGAYLGVVFGSVEGLVAVAVSLRVLFFVARPTTVLGINI	422	99US-0134370P.
Qy	525	PNSVIVRNVEHYQNAKHVPGMLILEIDAPIYFANASVLRERITRWIDEEERIKATGETS	584	99US-0134376P.
Db	423	PGTWVRMDQYAAAOQVPGVLRVDAPVYFANASYLRERSRWIDDEERTKSGEMG	482	99US-0134941P.
Qy	585	LQYIIDMGAIGNIDTSGISMLBEVKKITERRELQQLVLPVSVVMKKLNKSKFQNHGK	644	99US-0135124P.
Db	483	VRYVLDMGAIGSIDTSGTSMDELNKSILDRRCMQIVLANPGSEIMKKLDSKVLQIGH	542	99US-0135353P.
Qy	645	KWLYLTVEEAVGACNPNLRASKTNPKKD	672	99US-0136021P.
Db	543	EWVFPTGEAVASCDYVLVHSHKFGMAKD	570	99US-0136392P.
RESULT 6				
AAG32201				
XX	ID	AAG32201 standard; protein; 510 AA.		99US-0136782P.
XX	AC	AAG32201;		99US-0137222P.
XX	DT	17-OCT-2000 (first entry)		99US-0137528P.
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 38801.		99US-0137724P.
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;		99US-0138094P.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;			99US-0138540P.
KW	termination sequence.			99US-0138847P.
XX	OS	Arabidopsis thaliana.		99US-0139119P.
XX	PN	EP1033405-A2.		99US-0139452P.
XX	PD	06-SEP-2000.		99US-0139453P.
XX	PF	25-FEB-2000; 2000EP-00301439.		99US-0139454P.
XX	PR	25-FEB-1999; 99US-0121825P.		99US-0139455P.
PR	05-MAR-1999; 99US-0123180P.			99US-0139461P.
PR	09-MAR-1999; 99US-0123548P.			99US-0139462P.
PR	23-MAR-1999; 99US-0125788P.			99US-0139463P.
PR	29-MAR-1999; 99US-0126264P.			99US-0139457P.
PR	01-APR-1999; 99US-0127462P.			99US-0139458P.
PR	06-APR-1999; 99US-0128234P.			99US-0139459P.
PR	08-APR-1999; 99US-0128714P.			99US-0139460P.
PR	16-APR-1999; 99US-0129845P.			99US-0139461P.
PR	19-APR-1999; 99US-0130077P.			99US-0139462P.
PR	21-APR-1999; 99US-0130449P.			99US-0139463P.
PR	23-APR-1999; 99US-0130510P.			99US-0139464P.
PR	28-APR-1999; 99US-0130891P.			99US-0139465P.
PR	30-APR-1999; 99US-0131449P.			99US-0139466P.
PR	30-APR-1999; 99US-0132048P.			99US-0139467P.
PR	04-MAY-1999; 99US-0132407P.			99US-0139468P.
PR				99US-0139469P.

PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.
PR	03-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	03-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	14-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	23-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	23-SEP-1999;	99US-0155113P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-01583169P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	18-OCT-1999;	99US-0159638P.			
PR	21-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P			

XX	EP1033405-A2.	PR	01-JUL-1999;	99US-0141842P.
PN		PR	01-JUL-1999;	99US-0142154P.
XX		PR	02-JUL-1999;	99US-0142055P.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390P.
XX		PR	08-JUL-1999;	99US-0142803P.
PF		PR	09-JUL-1999;	99US-0142920P.
XX		PR	12-JUL-1999;	99US-0142977P.
XX	25-FEB-2000; 2000EP-00301439.	PR	13-JUL-1999;	99US-0143542P.
PR		PR	14-JUL-1999;	99US-0143624P.
PR	25-FEB-1999;	PR	15-JUL-1999;	99US-0144005P.
PR	05-MAR-1999;	PR	16-JUL-1999;	99US-0144085P.
PR	03-MAR-1999;	PR	16-JUL-1999;	99US-0144086P.
PR	23-MAR-1999;	PR	19-JUL-1999;	99US-0144325P.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144331P.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144332P.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144333P.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144334P.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144335P.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144352P.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352P.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192P.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151438P.
PR	24-JUN-1999;	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	PR	15-SEP-1999;	99US-0154018P.

```
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 59.5%; Score 2067; DB 3; Length 510;
Best Local Similarity 76.5%; Pred. No. 8.Se-188;
Matches 391; Conservative 65; Mismatches 53; Indels 2; Gaps 2;

QY 171 MLSNADVPNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLGHATTIIGFMGGAATV 230
DB 1 MLSKEVDAEKDKLYLHLAFTATLFAGVLEASLGIFRLGFIVDFLGHATTIIGFMGGAATL 60

QY 231 VCLQQLKSILGLBHFTHGADIISVMRSVFTQTHEWRWESAVLGCVPFIFLLSTRYFSKRR 290
DB 61 VSLQQLKGFGLKHTDSTDVISMRSVFSQTHEWRWESGVLCGFLFLLSTRYFSIKK 120

QY 291 PRFFWTSAMAPLTSVILGSLVLYFTAEKHGVVIGELKGLNPPSLTNLVFVSPYMTTA 350
DB 121 PKFFVAAAMAPLTSVILGSLVLYFTAEKHGVVIGELKGLNPLSGSLDIFTSPYMTSTA 180

QY 351 VKTGVVGIISLAEGTAGVSGFAMYNKYNIDGNKEMIAITGMNVGSGFTSCYLLTGPFSR 410
DB 181 VKTGLITGIITAAEGTAGVSGFAMYNKYNIDGNKEMIAFGMMNIVGSGFTSCYLLTGPFSR 240

QY 411 SAVYNAGCKTAASNIIMSLAVMLTLFTPLPHYTPLVLSAIIIVSAMGLGIDYEAH 470
DB 241 SAVYNAGCKTAMSNIVMAIVMFTLLFTPLPHYTPLVLSAIIISAMGLGIDYQAIIH 300

QY 471 LFKVDKDFVCMASVIGVVGVSVEIGLVIAIVISVLRLFTARPRFTVLGNIPNSVIY 530
DB 301 LMKVDKDFVCMASVIGVVGVSVEIGLVVAIVASIAARLLFVSRPKTAVKGNIPNSMIY 360

QY 531 RNVEHYONAKHVPCLILETDAPYFANASYLPERITRMTDEEEERIKATGETSLOYVII 590
DB 361 RNTEQYPPSSRTVPGLIILETDAPYFANASYLPERITRMTDEEEERIKATGETSLOYVII 420

QY 591 DMSAVGNIDTSGISMLEEYKKITERRELQVLVNPVSEVMKKLNKSKF-QNHILGKKWIYL 649
DB 421 DMSAVGNIDTSGISMVEIKKVIDRRALKVLVSNPKGEVVKLTKRSKFIGDHLGKEMFEL 480

QY 650 TYEEAVGACNFNLASKTNPKKDETEGMNV 680
DB 481 TVGEAVEACSYMLHTFKTEP-ASKNEPMNV 510

RESULT 8
AAY44942
ID AAY44942 standard; protein; 621 AA.
XX
AC AAY44942;
XX
DT 23-MAY-2000 (first entry)
XX
DE Soybean sulphate permease-1.
XX
KW Sulphate Permease; sulphate assimilation protein; soybean; probe;
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
KW antibody; screen.
XX
OS Glycine max.
XX
PN WO200004154-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015810.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Falco SC, Thorpe CJ;
XX
DR WPI; 2000-195025/17.
DR N-PSDB; AA250488.
XX
PT Nucleic acid fragments encoding sulfate assimilation proteins in plants
PT and seeds useful as probes for isolating cDNAs and genes encoding
PT homologous proteins, in producing transgenic plants.
XX
PS Claim 6; Page 50-52; 79pp; English.
XX
CC The present amino acid sequence is the soybean sulphate permease, a
CC sulphate assimilation protein. This sequence is obtained from ses2w
CC library, clone ses2w.pk0031.b3, derived from soybean embryogenic
CC suspension 2 weeks after subculture. It has 61% sequence identity to
CC Arabidopsis thaliana sulphate permease (gi 4579913). This sequence is
CC used as a probe to isolate other plant sulphate assimilation proteins,
CC for genetic and physical mapping of related genes and as markers of
CC traits linked to the gene. This is useful for plant breeding and to
CC construct chimeric genes, used to create transgenic plants with altered
CC levels of sulphate permease. The sulphate permease peptides are useful
CC for producing antibodies, that are used to screen and isolate cDNA clones
XX
SQ Sequence 621 AA;

Query Match 54.7%; Score 1901; DB 3; Length 621;
Best Local Similarity 59.0%; Pred. No. 7.7e-172;
Matches 359; Conservative 106; Mismatches 143; Indels 0; Gaps 0;

QY 48 PPQPFKSLKYSKETFPPDDPLRPFKNPKASKKMLGLQFPFPFEWAPKYTFQFLKAD 107
DB 3 PHQTTLKLRHVSEIFFPDDPLRPFKNQTRFKKLLALQYLFIPFDWAPNLTLLRSD 62

QY 108 LIAGTTIASLAIPQGISYAKLANLPILGLYSFTPLIYAMMGSSRDVAGTVAGSLL 167
DB 63 LISGLTIASLAIPQGISYAKLANLPILGLYSFTPLIYSLGSSRHVGVPVSLV 122
```


XX AAG32202; PR 18-JUN-1999; 99US-0139462P.
AC 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
DT 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
DE 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
KW 23-JUN-1999; 99US-0140354P.
KW 24-JUN-1999; 99US-0140695P.
KW 24-JUN-1999; 99US-0140823P.
XX 28-JUN-1999; 99US-0140899P.
OS 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
PN 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
PD 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
PF 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 18-JUN-1999; 99US-0139461P.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 38802.
XX Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134321P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137228P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.

QY	464	DYEAAILHLFKVDKDFVVCMSAYIGVFGSVEIGLVIAIVISVLRLVLLFIARPTFTVLGN	523
DB	241	DYQAAIHLWKVDKDFLVCMSAYVGVFGSVEIGLVVAISIAIARLLLFVSRPKTAVKGN	300
QY	524	IPNSVIYRNVEHYONAKHVPGLMLILEIDAPIYFANASYLRRITRWIDEERIKATGET	583
DB	301	IPNSMIYRNTEQYPPSSRTVPGLLILEIDAPIYFANASYLRRIRIWIDEERVKQSGES	360
QY	584	SLQYVIIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKKLNKSF-QNHIL	642
DB	361	SLQYIILDSAVGNIDTSGISMMVEIKKVIDRRALKVLVSNPKGEVVKLTSRKFIDHIL	420
QY	643	GKKWIYLTVEEAVGACFNFLRASKTNPKKDETEGWNV	680
DB	421	GKWMFLTVGEAVEACSYMLHTFKTEP-ASKNEPMNV	457
RESULT 11			
AA44935			
ID	RAY44935	standard; protein; 688 AA.	
XX	AC	RAY44935;	
XX	DT	23-MAY-2000 (first entry)	
XX	DE	Corn sulphate permease-1.	
XX	KW	Sulphate Permease; sulphate assimilation protein; corn; contig; probe;	
XX	KW	mapping; marker; plant breeding; chimeric gene; transgenic plant;	
XX	KW	antibody; screen.	
XX	OS	Zea mays.	
XX	PN	WO200004154-A2.	
XX	PD	27-JAN-2000.	
XX	PF	13-JUL-1999; 99WO-US015810.	
XX	PR	14-JUL-1998; 98US-0092833P.	
XX	PA	(DUPO) DU PONT DE NEMOURS & CO E I.	
XX	PI	Allen SM, Falco SC, Thorpe CJ;	
XX	DR	WPI; 2000-195025/17.	
XX	DR	N-PSDB; AA250481.	
XX	PT	Nucleic acid fragments encoding sulfate assimilation proteins in plants	
XX	PT	and seeds useful as probes for isolating cDNAs and genes encoding	
XX	PT	homologous proteins, in producing transgenic plants.	
XX	PS	Claim 6; Page 39-41; 79pp; English.	
XX	CC	The present amino acid sequence is the corn sulphate permease, a sulphate	
XX	CC	assimilation protein. This sequence is from a contig composed of clones	
XX	CC	cbn10.pk0062.b10, ccbln.pk081.h21, ccbln.pk092.12, cscic.pk005.j3,	
XX	CC	p0004.cblej58r, p0089.csdch19r, p0094.csesg12r, p0121.cfrmx30r and	
XX	CC	p0128.cbicz09r from cbn10, ccbln, cscic, p0004, p0089, p0094, p0121 and	
XX	CC	p0128 libraries, respectively. It has 80.2% sequence identity to	
XX	CC	Sporobolus stapifianus sulphate permease (gi 1907270). This sequence is	
XX	CC	used as a probe to isolate other plant sulphate assimilation proteins,	
XX	CC	for genetic and physical mapping of related genes and as markers of	
XX	CC	traits linked to the gene. This is useful for plant breeding and to	
XX	CC	construct chimeric genes, used to create transgenic plants with altered	
XX	CC	levels of sulphate permease. The sulphate permease peptides are useful	
XX	CC	for producing antibodies, that are used to screen and isolate cDNA clones	
XX	QY	Sequence 688 AA;	
Query Match 53.1%; Score 1843.5; DB 3; Length 688;			
Best Local Similarity 56.2%; Pred. No. 2.7e-166;			

Matches 354; Conservative 115; Mismatches 156; Indels 5; Gaps 3;	
Qy 41 VHQVEVPPQPEFFKSLKYSLKETFFDDPLRQFKNKPASKKFMGLQGFPPFFIFEWAPKYT 100	
Db 63 VHKVAPPARSTASKMKVRVKETFFDDPPQPFRAFQKQPGTQWLMAVRVLPFLDWDVPSYS 122	
Qy 101 FQFLKADLAIAGTITASLAIPQGISYAKLANUPPIILGLYSSFPPIPLIYMMGSSRDLAVGT 160	
Db 123 LSLFKSDLVAGLTIASLAIPQGISYAKLASLPPITGLYSSFPVPMVAVLGSSRDLAVGP 182	
Qy 161 VAVGSLMGMSLSNAVDNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLSHATI 220	
Db 183 VSISLLIINGMLRQAVSPFAETFLQLAFTSTLFPAGUQLASGLRLGRFVFLSKATL 242	
Qy 221 IGFMGGAATVVCILQKLSILGLEHETHGADIISVMRSVFTOTHEWRWESAVLGCVFIFFL 280	
Db 243 VGFMAAAIIVALQQLKGLLGIVHETTEMGIVPMASVFHHTSWSWGTLIMGVCLVFL 302	
Qy 281 LSTRYSKKRPFRFFVWSAMAPLTSVLGSLLYVFFTHAEKHGVEVIGELKKGLNPPSLTNL 340	
Db 303 LSAHVSIIRWPKLFWYSCAPLASVTISTLLVFLPKAQNHGISIIGLCKGLNRPSSWDKL 362	
Qy 341 VFVSPYMTTAVKTGIVVGGIISLAEGIAVGRSFAMKYKNYNIDGNKEMTAIGTMNVVGSFTS 400	
Db 363 LFDATYLGTMKTGLVTGIIISLATEGIAVGRTPASLKQYQIDGNKEMMAIGLNMVVGSGTS 422	
Qy 401 CYLTTGPPFSRSVAVNYNAGCKTAASNIIMSLAVMLTLLFLTPFLPHYTPLVLVLSAIIVSAML 460	
Db 423 CVYTTGAFSRSVAVNHAGCKTAMSNVIMALTVMVTLFLMLFPVFTPNVVLGALIIIAAVI 482	
Qy 461 GLIDYBAAIHLFKVDKDFVWCMSAYIGVFGSVEIGLIVIAIVSVLRVLLFTARPRTFV 520	
Db 483 GLIDPPAVYHIWKMDKMDFLVCVCAFAGVIFISVQEGLAIAVGISIFRVLVMOITRPRKMV 542	
Qy 521 LGNIPNSVIYRNVHAYQNAKHVPGMILLEIDAPIYFANASYLRRIETRWIDEBEERIKAT 580	
Db 543 QGNIKGTDIYRDLHHYKEAQRVSGFLIIAIEAPINFANSNLYNRIIRKWI--BEESFEQD 600	
Qy 581 GETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNK-SKFQ 639	
Db 601 KHTELHFIILDSAVPAIDTSGIAFLIIDIKSIKRGLELVLPVNPTEGEVWEKIQORANEAE 660	
Qy 640 NHLGKKWIIYLTVEEAVGACNFNLASKTNP 669	
Db 661 NYFRPDCLYLTGGEAI--ASLSALAKMTKP 688	
RESULT 12	
ABP53777	
ABP53777 standard; protein; 656 AA.	
AC ABP53777;	
XX DT 03-JAN-2003 (first entry)	
XX Arabidopsis thaliana sulphate transporter Sultr1;3 SEQ ID NO:2.	
DE Arabidopsis thaliana; sulphate transporter; chromosome 1; plant;	
XX sulphur-containing metabolite; glutathione; phytokeatin.	
XX Arabidopsis thaliana.	
OS JP2002272472-A.	
XX 24-SEP-2002.	
XX 22-MAR-2001; 2001JP-00082891.	
XX 22-MAR-2001; 2001JP-00082891.	
PR (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.	
PA WPI; 2002-718707/78.	
XX	
DR N-PSDB; AB082719.	
XX A gene encoding sulfate ion transporter Sultr1;3, a polynucleotide, an oligonucleotide, a recombinant vector, a plant.	
PT Claim 1; Page 9-10; 12pp; Japanese.	
XX The present invention describes a gene which is present in the first chromosome of Arabidopsis thaliana and encodes sulphate ion transporter Sultr1;3 having the present amino acid sequence. Also described: a polynucleotide purified from the genomic DNA, the mRNA, the cDNA or their complementary sequences; an oligonucleotide hybridising with the above gene or the above polynucleotide and consisting of a base sequence of at least 10 base pairs; a recombinant vector carrying the above polynucleotide; and a plant in which the above gene is introduced or its descendant or their tissues. The plant can be used for the preparation of sulphur-containing metabolites such as glutathione and phytokeatin	
XX	
SQ Sequence 656 AA;	
Query Match 52.0%; Score 1807.5; DB 5; Length 656;	
Best Local Similarity 53.7%; Pred. No. 6.9e-163;	
Matches 335; Conservative 119; Mismatches 169; Indels 1; Gaps 1;	
Qy 37 NFERVHQVEVPPQPFKSLKYSLKETFFDDPLRQFKNKPASKKFMGLQGFPPIFEWA 96	
Db 26 NTPYVHKVEVPPKQNLFNEFWYTFKETFFHDDPLRHFQDQSKKLMGIQSVFPVIEWG 85	
Qy 97 PKYTOFIKADLIIAGITIASLAIPQGISYAKLANLPIILGLYSSFPPIPLIYMMGSSRD 156	
Db 86 RYNNLKLFRGLIIAGLTIASLCPQDIGYAKLASLDPKYGLYSFVPLPYVACMGSSKDI 145	
Qy 157 AVGTAVGSLMGMSLSNAVDNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLS 216	
Db 146 AIGPVAVVSLGLTLLRABIDPNTNPNEYLRALFTSTFFAGVTQAAALGFRFLGLIDFLS 205	
Qy 217 HATIGFMGGAATVVCILQKLSILGLEHETHGADIISVMRSVFTOTHE--WRWESAVLGCV 275	
Db 206 HAAVVGFMGGAATIALQQLKGLGINKFTKTDIIAVLSVSISSAHGWNWQTILISAS 265	
Qy 276 FIFLLSTRYSKKRPFRFFVWSAMAPLTSVLGSLLYVFFTHAEKHGVEVIGELKKGLNPP 335	
Db 266 FLIFLLISKFIKGRNKKLFWIPAIAPLVSVIISTFFVYITRADKGGVQIVYKHLDKGLNPS 325	
Qy 336 SITNLVFSVPYMTTAVKTGIVVGGIISLAEGIAVGRSFAMKYKNYNIDGNKEMIAIGTMNVV 395	
Db 326 SURLYFSGDYLLKGFRIGVGVGVMVTEAVAGRTFAAMKDQIDGNKEMVAGMANVI 385	
Qy 396 GSFTSCYLTTGPPFSRSVAVNYNAGCKTAASNIIMSLAVMLTLLFLTPFLPHYTPLVLVLSAII 455	
Db 386 GSMTSCYVSTGFSRSVAVNFMAGQTAVSNIIMSIIVLLTLLFLTPFLKYTPNAILAAII 445	
Qy 456 VSAMGLIDYBAAIHLFKVDKDFVWCMSAYIGVFGSVEIGLIVIAIVSVLRVLLFIAR 515	
Db 446 INAVIPLVDVNATILIFKIDKLDFAVACMGAPFGVIFVSVEIGLLIIVAGISFAKILLQVTR 505	
Qy 516 PRTFVLGNIPNSVIYRNVHAYQNAKHVPGMILLEIDAPIYFANASYLRRIETRWIDEBEER 575	
Db 506 PRTAILGKPGTSVYRNINQYPEATRIPGVLTIRVDSAIYFNSNYSVYRRIQIQLWLTDEEE 565	
Qy 576 RIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNK 635	
Db 566 MVEAARLPRIQFLIEMSPVTDIDTSGIHALEDLYKSLQKEDQLVLANDPGPPVINKLHV 625	
Qy 636 SKFQNLGKKWIIYLTVEEAVGACN 659	
Db 626 SHFADLIGHDKIFLTVAEAVDSCS 649	
RESULT 13	
AAV4944	
ID AAV4944 standard; protein; 656 AA.	
XX	

[illegible]

PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142977P.
PR 12-JUL-1999; 99US-0143220P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 18-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 20-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 12-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2005, 08:53:14 ; Search time 33 Seconds
(without alignments)
1538.223 Million cell updates/sec

Title: US-10-762-049-18
Perfect score: 3475
Sequence: 1 HELARTLSYITHICLLRNTI.....NLRASKINPKKDETEGNNV 680

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3475	100.0	680	4	US-09-720-317A-18
2	2642.5	76.0	658	4	US-09-720-317A-24
3	2575.5	74.1	646	4	US-09-720-317A-26
4	2314	66.6	646	4	US-09-720-317A-28
5	2164	62.3	579	4	US-09-720-317A-4
6	1901	54.7	621	4	US-09-720-317A-16
7	1890.5	54.4	631	4	US-09-720-317A-29
8	1843.5	53.1	688	4	US-09-720-317A-2
9	1798.5	51.8	656	4	US-09-720-317A-20
10	1760.5	50.7	660	4	US-09-720-317A-30
11	1723.5	49.6	660	4	US-09-720-317A-23
12	1503	43.3	644	4	US-09-720-317A-25
13	997	28.7	685	4	US-09-720-317A-31
14	961.5	27.7	593	4	US-09-720-317A-22
15	731.5	21.1	228	4	US-09-720-317A-8
16	687.5	19.8	233	4	US-09-720-317A-27
17	687.5	19.8	595	4	US-09-252-991A-23812
18	675.5	19.4	311	4	US-09-720-317A-6
19	650.5	18.7	744	4	US-09-785-381-3
20	650	18.7	744	4	US-09-785-381-1
21	643.5	18.5	590	4	US-09-902-540-14944
22	641.5	18.5	748	4	US-09-949-016-10387
23	612	17.6	565	4	US-09-602-781A-616
24	587.5	16.9	780	4	US-09-785-381-11
25	586.5	16.9	764	1	US-08-424-567-2
26	586.5	16.9	764	2	US-08-711-928-2
27	586.5	16.9	764	3	US-09-184-937-2

28	586.5	16.9	790	4	US-09-949-016-11220	Sequence 11220, A
29	576.5	16.6	803	4	US-09-949-016-11498	Sequence 11498, A
30	567	16.3	828	4	US-09-248-796A-20746	Sequence 20746, A
31	549.5	15.8	663	4	US-09-875-811-6	Sequence 6, Appli
32	549	15.8	656	4	US-09-875-811-10	Sequence 10, Appli
33	546.5	15.7	679	4	US-09-875-811-2	Sequence 2, Appli
34	489	14.1	616	4	US-09-543-681A-4421	Sequence 4421, Ap
35	485.5	14.0	605	4	US-09-875-811-8	Sequence 8, Appli
36	485	14.0	598	4	US-09-875-811-12	Sequence 12, Appli
37	482.5	13.9	621	4	US-09-875-811-4	Sequence 4, Appli
38	454	13.1	970	4	US-09-795-927-7	Sequence 7, Appli
39	449	12.9	535	4	US-09-252-991A-21805	Sequence 21805, A
40	426.5	12.3	472	4	US-09-902-540-13374	Sequence 13374, A
41	402.5	11.6	575	4	US-09-438-185A-1015	Sequence 1015, Ap
42	381	11.0	159	4	US-09-720-317A-14	Sequence 14, Appli
43	369	10.6	566	4	US-09-543-681A-4544	Sequence 4544, Ap
44	363	10.4	533	4	US-09-107-532A-5244	Sequence 5244, Ap
45	360.5	10.4	147	4	US-09-720-317A-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1

US-09-720-317A-18
; Sequence 18, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Glycine max
US-09-720-317A-18

Query Match		100.0%;	Score 3475;	DB 4;	Length 680;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 680;		Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	HELARTLSYITHICLLRNTI	I	EDMGSDVDEYEP	PLGMNFRVHQVEVPPQPPFKSLKYSL 60
Db	1	HELARTLSYITHICLLRNTI	I	EDMGSDVDEYEP	PLGMNFRVHQVEVPPQPPFKSLKYSL 60
Qy	61	KETFPDDPLRQPKKPKSKMGLQFPF	F	FEWAPKYTEQFLKADLAGITIASLAIP 120	
Db	61	KETFPDDPLRQPKKPKSKMGLQFPF	F	FEWAPKYTEQFLKADLAGITIASLAIP 120	
Qy	121	QGISYAKLANLPILGLYSFIPPLIYAMGSSRD	L	AVGTAVGSLMGSMLSNAVDPNE 180	
Db	121	QGISYAKLANLPILGLYSFIPPLIYAMGSSRD	L	AVGTAVGSLMGSMLSNAVDPNE 180	
Qy	181	DPKLYLHLAFTATLFAGVFOAALGLFRLGLI	V	DFLSHATIIIGPMGAATVVCLOQLKSIL 240	
Db	181	DPKLYLHLAFTATLFAGVFOAALGLFRLGLI	V	DFLSHATIIIGPMGAATVVCLOQLKSIL 240	
Qy	241	GLEHFTHGADIIISVMRSVETQTHWRWESAV	L	GCVFIFLLSTRYFSKKRPRFFWVSAMA 300	
Db	241	GLEHFTHGADIIISVMRSVETQTHWRWESAV	L	GCVFIFLLSTRYFSKKRPRFFWVSAMA 300	
Qy	301	PLTSTVLGSLLYVFTHAEXHGVGEVIGELKK	L	GNLPPSLTNLVFVSPYMTTAVKTGIUVGII 360	
Db	301	PLTSTVLGSLLYVFTHAEXHGVGEVIGELKK	L	GNLPPSLTNLVFVSPYMTTAVKTGIUVGII 360	

Qy	361	SLAEGIAVGRSFAMYNQYNNIDGNKEMIAIGTNMNVVVGFTSCYLTGTGPFPSRSAVNNYAGCK	420
Db	361	SLAEGIAVGRSFAMYNQYNNIDGNKEMIAIGTNMNVVVGFTSCYLTGTGPFPSRSAVNNYAGCK	420
Qy	421	TAASNIIMSLAVMLTLLFTPLFHYTPVLVLSAIIVSAMLGLIDYEAAIHLFKVDKDFV	480
Db	421	TAASNIIMSLAVMLTLLFTPLFHYTPVLVLSAIIVSAMLGLIDYEAAIHLFKVDKDFV	480
Qy	481	VMSAYIGVVGVSVEIGLVIATIVISVLRLVLFIAARPTFVLGNIPNSVIYRNVHEYQNAK	540
Db	481	VMSAYIGVVGVSVEIGLVIATIVISVLRLVLFIAARPTFVLGNIPNSVIYRNVHEYQNAK	540
Qy	541	HVPGMLILBIDAPIYFANASYLRERITRWIDEEBERIKATGETSLQVVIIDMSAVGNIDT	600
Db	541	HVPGMLILBIDAPIYFANASYLRERITRWIDEEBERIKATGETSLQVVIIDMSAVGNIDT	600
Qy	601	SGISMLEEVKKITERRRELQVLVNPVSEVMKKLNKSKFQNHGLGKKWYLTVEEAVGACNP	660
Db	601	SGISMLEEVKKITERRRELQVLVNPVSEVMKKLNKSKFQNHGLGKKWYLTVEEAVGACNP	660
Qy	661	NLRASKTNPKKDETEGWNV	680
Db	661	NLRASKTNPKKDETEGWNV	680
RESULT 2			
US-09-720-317A-24			
; Sequence 24, Application US/09720317A			
; Patent No. 6696292			
; GENERAL INFORMATION:			
; APPLICANT: Stephen M. Allen			
; APPLICANT: Saverio C. Falco			
; APPLICANT: Catherine J. Thorpe			
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins			
; FILE REFERENCE: BB-1167			
; CURRENT APPLICATION NUMBER: US/09/720,317A			
; CURRENT FILING DATE: 2000-12-21			
; PRIOR APPLICATION NUMBER: 60/092,833			
; PRIOR FILING DATE: 14-07-1998			
; NUMBER OF SEQ ID NOS: 31			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 24			
; LENGTH: 658			
; TYPE: PRT			
; ORGANISM: Arabidopsis thaliana			
US-09-720-317A-24			
Query Match			
Best Local Similarity 76.0%; Score 2642.5; DB 4; Length 658;			
Matches 501; Conservative 80; Mismatches 75; Indels 3; Gaps 3			
Qy	24	MGSVDYEPGLMNNRERVHQ-VEVPPPOPFKSLKYSIKETFPDPDPLQFKNKPKASKP	82
Db	1	MGTEDTFPQGAEEHLRHHTVEAPQOPFLKSLQYSVKETLPDPPFPQFKNQNSRKP	60
Qy	83	MLGLOFFPPIFWAPKTYTQFLKADLAGITIASLAIPQGISYAKLANLPILGLYSSPI	142
Db	61	VLGLKYFLPIFWAPRYMLKFFKSDLINGITIASLAIPQGISYAKLANLPILGLYSSV	120
Qy	143	PPLIYAMGSSRDLAGVTGAVGSLGMSLMSNAVDPNEDPKLYHLHATATLFAGVQAA	202
Db	121	PPLVAVVLGSSRDLAGVTGAVVASULLTGAMLSKEVDAEKDPKLYHLHATATLFAGVLEAS	180
Qy	203	LGLFRLGLIVDFLSHATIIIPMGNGAATVVCLOQLKSLIGLHETHFHGADIIISVRSVFTOT	262
Db	181	LGIIFRLGPTVDFLSHATIIIPMGNGAATVVSLOQLKGIKGFHKTDTSDTIVISVRSVFSOT	240
Qy	263	HEWRVESAVLGCVFPIFELLSTRYSKKRPFVWSAMAPLTSVLGSLAVYFTHAEKHGV	322
Db	241	HEWRVESVGLGCGLFFLLSTRYSIKKPFVWAANAPLTSVLGSLAVYFTHAEKHGV	300
Qy	323	EVIGELKKGLNPPSLTNLFFVSPYMTTAVKTGIVVGIIISLAEGIAVGRSFAMYNKNIDG	382

Db 301 QVIGDLKKGLNPLGSDLIFTSPYMSTAVKTLGTIGTIIALAEGLAVGRSFAMFKNNYIDG 360
 Qy 383 NKEMIAIGTMNVVGSFTSCYLTGTPFSRSVNNYAGCKTAASNIIIMSLAVMLTLLELTPL 442
 Db 361 NKEMIAFGMMNIVGSFTSCYLTGTPFSRSVNNYAGCKTAMSNIVMAIVAMFTLLFLTPL 420
 Qy 443 PHYTPLVVLIAIIVSAMLGLIDYENAAIHLFKVDKPDFVVCMSAYIGVVPFSGVEIGLVIAI 502
 Db 421 PHYTPLVVLIAIISAMGLIDYQAAIHLWKVDKPDFLVCMSAYYGVVPFSGVEIGLVAV 480
 Qy 503 VISLVRLLFTARPTFVLGNIPNSVIYRNVHEYQNAKHVPQMLTLEIDAPIYFANASYL 562
 Db 481 AISIARLLLFVSRPKTAVKGNIPNSMIVRNTQEPSSRTVPGILILEIDAPIYFANASYL 540
 Qy 563 RERITRWIDEEERIKATGETSLOYVIIDMSAVGNIDTSGISMLBEVKKITERRELQVLV 622
 Db 541 RERIIRWIDEEERVKQSGESSLOYIILDM SAVGNIDTSGISMMVEIKKVIDRRALKLVL 600
 Qy 623 VNPVSEVMMKLNKSKP-QNHLGKWIYLTVEEAVGACNPLRASKTNPKDETEGWNV 680
 Db 601 SNPKGEVVKLTRSKFTGDHLGKEMFELTVGEAVEACSYMLHTFKTEP-ASKNEPWNV 658

 RESULT 3
 US-09-720-317A-26
 ; Sequence 26, Application US/09720317A
 ; Patent No. 6696292
 ; GENERAL INFORMATION:
 ; APPLICANT: Stephen M. Allen
 ; APPLICANT: Saverio C. Falco
 ; APPLICANT: Catherine J. Thorpe
 ; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
 ; FILE REFERENCE: BB-1167
 ; CURRENT APPLICATION NUMBER: US/09/720,317A
 ; CURRENT FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: 60/092,833
 ; PRIOR FILING DATE: 14-07-1998
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 26
 ; LENGTH: 646
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-720-317A-26

 Query Match 74.1%; Score 2575.5; DB 4; Length 646;
 Best Local Similarity 74.4%; Pred. No. 5e-251;
 Matches 490; Conservative 80; Mismatches 74; Indels 15; Gaps 4

 Qy 24 MGSVDYEPPLGNMNFVRHQ-VEVPPPPQPPFKSLKYSLKETFFPDPLRQPKNKPKASKKF 82
 Db 1 MGTEDTYFPQGAEBLRRHHHTVEAPQPQPFLLKLSQYSVKETLFPDPPRQPKNQNASRKF 60

 Qy 83 MLGQFPFPIPEWAPKPTFFQKADLAGITIASLAIPQGISYAKLANLPPILGLYSGFI 142
 Db 61 VLGLKYFLPIFEWAPRPNLKFKSDLAGITIASLAIPQGISYAKLANLPPILGLYSSVF 120

 Qy 143 PPLIYAMMGSSRDLAGVTAVAGSLLMGSMLSNAVDPNEDPKLYLHLAFTATLFAGVFOAA 202
 Db 121 PPLVYAVLGSSRDLAGVTAVASLLTGAMLSKEVDAEKDPKLYLHLAFTATFFAGVLEAS 180

 Qy 203 LGLFRLGLIVDFLSHATIGPMGAATVVCILQOLKSLILGLBHPFHGADIISVMRSVFQT 262
 Db 181 LGLFRLGLIVDFLSHATIGPMGAATVVSILQOLKGFGLKHFTDSTDVIVSMRSVFQT 240

 Qy 263 HEWSEWAGVLCGVIFIFLLSTRYSFKRPRFVWSAMAPLTSVILGSLLVYFTHAEKHGV 322
 Db 241 HEWSEWAGVLCGVIFLLSTRYSIFKKPKFVWAAAPLTSVILGSLLVYFTHAEHGV 300

 Qy 323 EVIGELKKGLNPPSLTNLVFVSPYMTAVKGTGVVGIISLAEGIAVGRSFAMFKNNYIDG 382
 Db 301 QV-----GSDLIFTSPYMSTAVKTLGTIGTIIALAEGLAVGRSFAMFKNNYIDG 348

QY 383 NKEMIAIGTMNVGSFTSCVLTTPGPFSSAVNNAGCKTAASNIIMSLAVMLTLLFLTPL 442
Db 349 NKEMIAFGMNI VGSFTSCVLTTPGPFSSAVNNAGCKTAMNI VMAIMFTLLFLTPL 408
QY 443 FHYTPLVLSAII VSMAGLIDYEAATHLKFVDKDFVVCMSAYIGVVGFSVEIGLVIAI 502
Db 409 FHYTPLVLSAII VSMAGLIDYEAATHLKFVDKDFVVCMSAYIGVVGFSVEIGLVIAI 468
QY 503 VISLRLVLLFIAPRTFVLGNIPNSVIYRNVEHYQNAKHVPGMLILLEIDAPIYFANASYL 562
Db 469 AISIARULLLVSRPKTAVKGNIPNSMIYRTEQYPPSSRTVPGILILEIDAPIYFANASYL 528
QY 563 RERITRWIDEEERIKATGETSLOYVIDMSAVGNIDTSGISMLEEVKKITERRELQVL 622
Db 529 RERIIRWIDEEERVKOGSSSLQYIILDSAVGNIDTSGISMMVEIKVIDRALKVL 588
QY 623 VNPVSEVMKKLNKSKP-QNHGLKKWYLYTVEEAVGACNFNLRAKTNPKKDETEGNNV 680
Db 589 SNPKGEVVKLTIRSKFTGDHLGKEMELTVGEAVEACSYMLHTFKTEP-ASKNEPNNV 646

RESULT 4
US-09-720-317A-28
; Sequence 28, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-720-317A-28

Query Match 66.6%; Score 2314; DB 4; Length 646;
Best Local Similarity 67.2%; Pred. No. 1.3e-224;
Matches 430; Conservative 104; Mismatches 102; Indels 4; Gaps 2;

QY 42 HOVEVPPQPFSLKYSKATETFPDDPLQFKNK-PASKKFWLGLQFFPIFEWAPKYT 100
Db 10 HOVEIPPPQPLSKNTLNEILFADDPFRIRNESKTSKKELGRHVFPFILEWARGYS 69
QY 101 FQFLKADLIAGITIASIAIPQGISYAKLANLPILGLYSSFIPLIYAMGSSRDLA VGT 160
Db 70 LEYKSDVISGITIASIAIPQGISYAKLANLPILGLYSSLPVLYAINGSSRDLA VGT 129
QY 161 VAGSLLMGMLSNVADPNEDPKLYLHLATATLFAGVFOAALGLPRLGLIIVDFLSHATI 220
Db 130 VAVASLLTAAMLGKEVNNAVNPVKLYLHLATATLFAGVFOAALGLPRLGLIIVDFLSHATI 189
QY 221 IGFMGGAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQTHEWRWESAVLCVFIPL 280
Db 190 VGFMGGAATVVCLOQLKGLGLHHFTHTDITVLRISFQSHRWESVGLGCCFLIFL 249
QY 281 LSTRYSKGRPRFVWSAMAPLTSVILGSLLVYFTTHAEKGVGEVIGELKKGLNPPSLTNL 340
Db 250 LTTKYISKRPKLFWISAMSPVSVIFGTIFLFLHDQFHGIQFICELKKGINPSTHIL 309
QY 341 VFVSPYMTTAVKGIIVVGIISLAEGIAVGRSFAMKYKNYIDGNKEMIAICTMNVGSFTS 400
Db 310 VFTPPYMLALKVGIITGVIALAEGIAVGRSFAMKYKNYIDGNKEMIAIFGMNIIIGSFSS 369
QY 401 CYLTTGPFSSAVNNAGCKTAASNIIMSLAVMLTLLFLTPLPHYPLVLSAII VSMAGL 460

Db 370 CYLTTGPFSSAVNNAGCKTALSNNVMAVAVAVTLLFLTPLFFYTPLVVLSIIIAAML 429
QY 461 GLIDYEAATHLKFVDKDFVVCMSAYIGVVGFSVEIGLVIAI VLSLRLVLLFIAPRTFV 520
Db 430 GLVDYEAATHLWKLDKDFDFVCLSVGLVGVFGTIEIGLILSVGISVMRLVLFVGRPKIYV 489
QY 521 LGNIPNSVIYRNVEHYQNAKHVPGMLILLEIDAPIYFANASYL RERITRWIDEEERIKAT 580
Db 490 MGNIQNSEIYRNIEHYPOAIRSSLLILHIDGPIYFANSTYLRDRIGRWIDEEEDKLRTS 549
QY 581 GETSLOYVIDMSAVGNIDTSGISMLEEVKKITERRELQVL VNPVSEVMKKLNKSKFQ 640
Db 550 GDISLQYIILDSAVGNIDTSGISMLEEELNKILGRRELKLVIANPGAEVMMKLSKSTFIE 609
QY 641 HLGKWIYLYTVEEAVGACNFNLRAKTNPKKDETEGNNV 680
Db 610 SIGKERIYLYTVAEVAACDFMLHTAKPDSVPPE---FNNV 646

RESULT 5
US-09-720-317A-4
; Sequence 4, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Zea mays
US-09-720-317A-4

Query Match 62.3%; Score 2164; DB 4; Length 579;
Best Local Similarity 71.0%; Pred. No. 1.6e-209;
Matches 403; Conservative 86; Mismatches 79; Indels 0; Gaps 0;

QY 105 KADLIAGITIASIAIPQGISYAKLANLPILGLYSSFIPLIYAMGSSRDLA VGTAVG 164
Db 3 ESDLIAGITIASIAIPQGISYAKLANLPVGLYSSFPVPLVYALMGSSKOLA VGTAVGA 62
QY 165 SLLMGMLSNVADPNEDPKLYLHLATATLFAGVFOAALGLPRLGLIIVDFLSHATIIGFM 224
Db 63 SLLISMLGSEVSPTEPNVLYLHLATATLFAGVFOAALGLPRLGLIIVDFLSHATIIGFM 122
QY 225 GGAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQTHEWRWESAVLCVFIPLSTR 284
Db 123 AGAATVVCLOQLKGLGLHHFTHTDITDVSVMSVFSQTHQWRWESVLLCGCFLLVTR 182
QY 285 YFSKGRPRFVWSAMAPLTSVILGSLLVYFTTHAEKGVGEVIGELKKGLNPPSLTNLVFVS 344
Db 183 FISKRPKLFWISAAAPLTSVILGSLVLVYLTHAENHGVIGVYLLKGLNPPSVTSLOFSP 242
QY 345 PYMTTAVKGIIVVGIISLAEGIAVGRSFAMKYKNYIDGNKEMIAICTMNVGSFTSCYL 404
Db 243 PYMMLALKGTIITGVIALAEGIAVGRSFAMKYKNYIDGNKEMIAICTMNVGLSLSYLT 302
QY 405 TGPFSRSVNNAGCKTAASNIIMSLAVMLTLLFLTPLPHYPLVLSAII VSMAGLID 464
Db 303 TGPFSRSVNNAGCKTAMSNVMSLAVNMVLTLLFLTPLPHYPLVLSAII VSMAGLVD 362
QY 465 YEAAIHLKFVDKDFVVCMSAYIGVVGFSVEIGLVIAI VLSLRLVLLFIAPRTFVLGNI 524
Db 363 FGNALHLWRVDKDFVCAGAVLGVVGSVEGLVAVASLRLVLLFVARPTTVLGN 422

QY 581 GETSLQYVVIIDMSAVGNIDTSGISMLBVEVKKITERRELQVLVNPVSEVMKKLNKSKFON 640
Db 543 KHSLSQLLIEMSAVSGVDNGVSFFKELKKTAKKDIELVFNPLSEVVVKLQRADEQK 602
QY 641 H-LGKWKIYLTVEAVGA 657
Db 603 EFMPEPFLJTVAEAVAS 620

RESULT 8
US-09-720-317A-2
; Sequence 2, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Zea mays
US-09-720-317A-2

Query Match 53.1%; Score 1843.5; DB 4; Length 688;
Best Local Similarity 56.2%; Pred. No. 5.2e-177;
Matches 354; Conservative 115; Mismatches 156; Indels 5; Gaps 3;

QY 41 VHOVEVPPQPPFFKSLKYSLKETFFDDPLRQPKPKPASKKFMGLQGFPPPIFEWAPKYT 100
Db 63 VHKVAPPARSTAKMKVRKETFFDDPPRAFKGQPGTQWLNAVRYLFPILDWPSVGS 122
QY 101 FQFLKADLIAGITIASLAIPOGISYAKLANLPILGLYSSFPPLIYAMGSSRDIAVGT 160
Db 123 LSLFKSDLVAGLTIASLAIPOGISYAKLASLPILGLYSSFPVPMVYAVLGSSRDIAVGP 182
QY 161 VAVGSLMGSMNSNAVDNPDNPKLYLHLAFTAILPAGVFOAALGLFRLGLIVDFLSHAT 220
Db 183 VSISLIMGSLRAQVSPTEPTLFLQLAFTSTLFAGLVQASGLILRLGFVIDFLSKATL 242
QY 221 IGFMGGAATVVCLOQLKSIILGLEHFTHGADIISVMSRVFTQTHEWRWESAVLGCVFIFL 280
Db 243 VGFMAAGAAITIALQQLKYVLGIANFTRKTDIVSVMSVRSVHHGWNMQTIVIGVSFLVF 302
QY 281 LSTRYFSKKRPRFPWVSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGLNPPSLTNL 340
Db 303 LSARHVSIRWPKLFWVSACAPLASVTISTLLVFLFKAQNHGIIIGQLKGLNRPSSDKL 362
QY 341 VFVSPYMTTAVKTVGVVGIISLAEGIAVGRSPFAMYKNYNIDGNKEMIAIGTMNVVGSFTS 400
Db 363 LFDATVILGLTMTKGLTVGTIISLATEGIAVGRFTASLDKYDIDGNKEMWIAIGLMNVVGSCTS 422
QY 401 CYLTTGPPFSRANVYAGCKTAASNIIMSLAVMTLLFLTPLFHYTPLVLSAIIYSAML 460
Db 423 CYVTGTAFPSRANVHAGCKTASNSVIMALTVMVTLFLFPLFVYTPNVVILGAIITAAVI 482
QY 461 GLIDYEAAILHFKVDKDFVVCMSAYIGVVGVSVEIGLVIAIVISVLRLVLLFTIARPTFV 520
Db 483 GLIDFPVAVHIWKMDKMDFLVCVCAFAGVIFISVQEGLAIVAGISIFRVLQMQITRPMV 542
QY 521 LGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRRERTRWIDEEERIKA 580
Db 543 QGNIKGTDIYRDLHHYKEAQRVSGFLTILATEAPINFANSYLNRIKRWI--EESFEQD 600
QY 581 GETSLQYVVIIDMSAVGNIDTSGISMLBVEVKKITERRELQVLVNPVSEVMKKLNK-SKFQ 639

Db 601 KHTELHFIILDSAVPAIDTSGIAFLIDIKSIEKGLGLVLVNPVGEVMEKIQRANEAE 660
QY 640 NHLGKWKIYLTVEEAVGACNFNLRAKTNP 669
Db 661 NYFRPDCLYLTGGEAI--ASLSALAKWTKP 688

RESULT 9
US-09-720-317A-20
; Sequence 20, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-720-317A-20

Query Match 51.8%; Score 1798.5; DB 4; Length 656;
Best Local Similarity 53.6%; Pred. No. 1.7e-172;
Matches 332; Conservative 123; Mismatches 163; Indels 1; Gaps 1;

QY 41 VHOVEVPPQPPFFKSLKYSLKETFFDDPLRQPKPKPASKKFMGLQGFPPPIFEWAPKYT 100
Db 30 VYKVGYPKPNLATEFTETLRSTFFHDNPLRQYKQSGPRRFRMGLFLFPIFGWGRDYS 89
QY 101 FQFLKADLIAGITIASLAIPOGISYAKLANLPILGLYSSFPPLIYAMGSSRDIAVGT 160
Db 90 LNKFKGDLTAGLTIASLAIPOGISYAKLANLPQGLYSFIPPLIYAAAGSSRDIAIGP 149
QY 161 VAVGSLMGSMNSNAVDNPDNPKLYLHLAFTAILPAGVFOAALGLFRLGLIVDFLSHAT 220
Db 150 VAVVSLIIGSLLOAEVDHVNKEEYMLRAFTATFFAGITQAAALGLFRLGLFLFLSHAAI 209
QY 221 IGFMGGAATVVCLOQLKSIILGLEHFTHGADIISVMSRVFTQTHE-WRWESAVLGCVFIF 279
Db 210 VGFMGGAATIALQQLKYVLGIANFTRKTDIVSVMSVRSVHHGWNMQTIVIGVSFLVF 269
QY 280 LSTRYFSKKRPRFPWVSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGLNPPSLTN 339
Db 270 LLFAKYIGKKRKLFWVPAIAPISVILATFFVYITRADKQGVQIVKHIEQGINPSSVHK 329
QY 340 LVFVSPYMTTAVKTVGVVGIISLAEGIAVGRSPFAMYKNYNIDGNKEMIAIGTMNVVGSFT 399
Db 330 IYFTGPFVAKGPKIGVVGCVIGLTEAVAIGRTFAAMQYQLDGNKEMVALGTNIVGSM 389
QY 400 SCYLTGPPFSRANVYAGCKTAASNIIMSLAVMTLLFLTPLFHYTPLVLSAIIYSAM 459
Db 390 SCYVTTGFSRANVFMACCKTPVSNVMSVVLLTLVITPLFKYTPNAILGSIISAV 449
QY 460 LGIDYEAAILHFKVDKDFVVCMSAYIGVVGVSVEIGLVIAIVISVLRLVLLFIARPTF 519
Db 450 IGLVDYEAAILWKVDKDLDFIACMGAFVGVFVSEIGLIIAIVISFAKILLQVTRPTA 509
QY 520 VLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRRERTRWIDEEERIKA 579
Db 510 LIGNPGTTIYRNISQYPEAKLTGPGVIVRVDSAIYFSNSVYRERILLRWLTDSEERAKA 569
QY 580 TGSETSQQYVVIIDMSAVGNIDTSGISMLBVEVKKITERRELQVLVNPVSEVMKKLNKSKFQ 639
Db 570 VGLPKTSFLIVEMSPVIDIDTSGIHALEDLYKNLQKKMQILISNPGSVVIEKIQASKLT 629

RESULT 12

US-09-720-317A-25
; Sequence 25, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Stylisanthes hamata
US-09-720-317A-25

Query Match 43.3%; Score 1503; DB 4; Length 644;
Best Local Similarity 47.0%; Pred. No. 1.2e-142; Mismatches 124; Indels 22; Gaps 6;
Matches 301; Conservative 124; Mismatches 193; Indels 22; Gaps 6;
QY 33 LGMNFRVHQ--VEVPPPPQPFKSLKYSKETFPPDDPLRQFNKPKASKKFMGLQF-- 88
DB 4 LGTEQFSERQWLNPNPLTKKFLGPLK-----NKFTSSSSKKEKTRVNSFLA 55
QY 89 -FPPIFWAPKYTOFKADLIAGITIASIAPOGISYAKLANLPPIILGYSGFIPPLIY 147
DB 56 SLFPILSWIRTSATKFKDGLSLGLTLASLSIPQSIGYANLAKLDPOGLYTSVIPVIV 115
QY 148 AMGSSROLAVGTAVGSLMGMLSNVDPNEDPKLYLHLATATLPAQVQAAQLGR 207
DB 116 ALMGSSREIAGPVAVVVMSSLSLVKVIDPDAPHNDYRLNLFVTTLFAGIFOTAFGLR 175
QY 208 LGLIVDFLSHATITGFMGGGAATVVCLOQLKSLGLEHFTHGADIISVMRSVFTQTHE-- 264
DB 176 LGFLVDFLSHAALVGFAGAAVIGLQQLKGLLGLTHFTTKTDAVALKSVYTSLHQIIT 235
QY 265 ----WRWESAVLGCVFIFILLSTRYFSKRPFRFWSAMAPLTSVILGSLLVYFTHAEKH 320
DB 236 SSENWSPNLVFGCSFLIFLLAARFICRRNKKFFWLPALAPLLSVILSTLIVFLSKGDKH 295
QY 321 GVEVIGELKGLNPPSLTNLVFVSPYMTTAVKTIIVGIIISLAEGIAVGRSFAMKYNYI 380
DB 296 GWNIIKHVQGLNPSVHKLQNLGPHVGOAKIGLSAIIATLTAIAVGRSPANIKGYHL 355
QY 381 DGNKEMIAIGTMNVVSGFTSCYLTGTPFSAVNYNAGCKTAASNIIMSLAVMLTLFLT 440
DB 356 DGNKEMLAMGCMNIAGSLTSCYSTGTSFTAVNFSAGCKTAVSNIVMAYTVLLCLELFT 415
QY 441 PLPHYTLVLSIAIVSAMLGLIDYEAAILFKVDKDFVVCMSAYIGVVGVSGVEIGLVI 500
DB 416 RLLVYTPMAILASIIISALPGLIDIGEAVHWKVDKDFLACLGAPFGLVFSIEIGLLI 475
QY 501 AIVISVLVLLFIARPTFLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDA-PIYFANA 559
DB 476 ALSISPAKILLQAIIRPGVGLGRIPPTTEAYCDVAQYPMAYVTPGILVIRISSGLCFANA 535
QY 560 SYLRERITRWI-DEEBERIKATGTSLOQYIIMDSAVGNIDTSGISMLEBVKKITERREL 618
DB 536 GFVRERILKWEDEQNIIEAAKGRVQAIIDMTDLTNDTSGILALELHKLLSRGV 595
QY 619 QLVLVNPSVEMKXKLNKSKFQNLGKKWIVLTVVEAVGAC 658
DB 596 ELAMVNRPRWEIHKLVANFVDKIGKERVFLTVAEAVDAC 635

RESULT 13

US-09-720-317A-22
; Sequence 22, Application US/09720317A

US-09-720-317A-31
; Sequence 31, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-720-317A-31

Query Match 28.7%; Score 997; DB 4; Length 685;
Best Local Similarity 34.0%; Pred. No. 1.8e-91;
Matches 217; Conservative 140; Mismatches 232; Indels 50; Gaps 13;
QY 47 PPOQPFKSLKYSKETFPPDDPLRQFNKPKASKKFMGLQFPPFIFEWAPKYTF-QFLK 105
DB 51 PPSIPE-----DDIFSGWTAKIKRMRLVDWIDTLFPCPRWIRTYRWSEYFK 96
QY 106 ADLIAGITIASIAPOGISYAKLANLPPIILGYSSIPPLIYAMGSSRDLAVGTAVGS 165
DB 97 LDMAGITIGVIMLVPOAMSAYAKLAGLPPYIGLYSSFPVYVYAIAGSSRQLAIGPVALVS 156
QY 166 LMGSMLSNAVDNEDPKLYLHLATATLPAQVQAAQLGRFLGLIVDFLSHATITGFMG 225
DB 157 LLVSNALGGIADNTEE--LHIELAILLALLVGLIEGMLLGLRLGLWIRFISHSVISGFTS 214
QY 226 GAATVVCLOQLKSLGLEHFTHGADIISVMRSVFTQTHEWRWESAVLGCVFIFILLSTRY 285
DB 215 ASAVITGLSIIKYLGLYS-IARSSKIVPIVESIIAGADKQFQWPPFVMSGLILVILQVMKH 273
QY 286 FSKRPRFRFWSAMAPLTSVILGSLLVYFTHAEKHGVEVIGELKGLN----PPLTNLV 341
DB 274 VGKAKKELQFLRAAAPLTVGLTIAKVPHPP--SISLVGEIPOGLTFTSFPRSPDHAK 331
QY 342 FVSPYMTTAVKTIIVGIIISLAEGIAVGRSFAMKYNYNIDGNKEMIAIGTMNVVSGFTSC 401
DB 332 TLLP--TSALITG-----VPILSVGIKALAAKNRYELDSNLDLFGLVANILGSLFSA 384
QY 402 YLTGTPFSAVNYNAGCKTAASNIIMSLAVMLTLFLTPLFHYTFLVLSIAIVSAMLG 461
DB 385 YPATGSFSAVNNESEAKTGLSLITGIIIGCSLLFLTPMPKYIPOCALAAIVISAVSG 444
QY 462 LIDYEAAILFKVDKDFVVCMSAYIGVVGVSGVEIGLVIIVISVLVLLFIARPTFV 520
DB 445 LDYDIAIFLWRDKRDFSLWTIT 503
QY 521 LGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRERITRW----IDE----- 572
DB 504 LGRPLGTTVYRNIIKQYPEAYTYNGIVIVRIDSPIYFANISYIKDRLEVEVAVDKYNRG 563
QY 573 -EERIKATGTSLOQYIIMDSAVGNIDTSGISMLEBVKKITERRELQVLVNPVSEVMK 631
DB 564 LEVDRI-----NFVILEMSPVTHIDSSAVEALKELYQYKTRDITQLAISPNKDVHL 615
QY 632 KLNKSKFQNLGKKWIVLTVVEAVGACNPLRASKTNPK 670
DB 616 TIARSGMVELGKWFVFRVHDVAQVQCLOVVOSSNLEDK 654

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2005, 09:39:29 ; Search time 56 Seconds
(without alignments)
4771.247 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLRNTI.....NLRASTNPKDTEGNNV 680

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpa/US11_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2330.5	67.1	657	US-10-437-963-140667	Sequence 140667,
2	2015	58.0	414	US-10-424-599-273958	Sequence 273958,
3	1853.5	53.3	695	US-10-425-115-235264	Sequence 235264,
4	1850	53.2	659	US-10-424-599-207945	Sequence 207945,
5	1811.5	52.1	666	US-10-437-963-198336	Sequence 198336,
6	1773.5	51.0	662	US-10-437-963-179488	Sequence 179488,
7	1762	50.7	681	US-10-425-115-260426	Sequence 260426,
8	1733.5	49.9	681	US-10-437-963-175893	Sequence 175893,
9	1724	49.6	397	US-10-424-599-273959	Sequence 273959,
10	1723.5	49.6	653	US-10-437-963-179495	Sequence 179495,
11	1722.5	49.6	662	US-10-425-115-366870	Sequence 366870,

12	1563.5	45.0	659	US-10-437-963-174545	Sequence 174545,
13	1500	43.2	653	US-10-424-599-195017	Sequence 195017,
14	1460	42.0	575	US-10-425-114-69272	Sequence 69272, A
15	1436	41.3	405	US-10-425-115-228851	Sequence 228851,
16	1406	40.5	640	US-10-425-115-224859	Sequence 224859,
17	1379.5	39.7	687	US-10-424-599-217235	Sequence 217235,
18	1344.5	38.7	587	US-10-437-963-180840	Sequence 180840,
19	1222	35.2	490	US-10-424-599-225979	Sequence 225979,
20	1197.5	34.5	730	US-10-437-963-178936	Sequence 178936,
21	1185.5	34.1	543	US-10-424-599-217231	Sequence 217231,
22	1018.5	29.3	330	US-10-424-599-282576	Sequence 282576,
23	1013	29.2	332	US-10-424-599-197192	Sequence 197192,
24	1011.5	29.1	339	US-10-424-599-273793	Sequence 273793,
25	998	28.7	485	US-10-739-930-9381	Sequence 9381, Ap
26	994	28.6	485	US-10-424-599-203123	Sequence 203123,
27	975.5	28.1	392	US-10-425-115-260428	Sequence 260428,
28	955.5	27.5	689	US-10-437-963-195183	Sequence 195183,
29	923	26.6	271	US-10-767-701-43225	Sequence 43225, A
30	921	26.5	224	US-10-767-701-32265	Sequence 32265, A
31	905	26.0	323	US-10-767-701-40999	Sequence 40999, A
32	887	25.5	303	US-10-437-963-188282	Sequence 188282,
33	833.5	24.0	522	US-10-437-963-109890	Sequence 109890,
34	794	22.8	156	US-10-424-599-175150	Sequence 175150,
35	784.5	22.6	187	US-10-424-599-175149	Sequence 175149,
36	737.5	21.2	279	US-10-424-599-161259	Sequence 161259,
37	730.5	21.0	280	US-10-424-599-278772	Sequence 278772,
38	715	20.6	262	US-10-425-115-260430	Sequence 260430,
39	697	20.1	265	US-10-424-599-246655	Sequence 246655,
40	674	19.4	227	US-10-425-115-366867	Sequence 366867,
41	668	19.2	714	US-09-749-589-4	Sequence 4, Appli
42	668	19.2	714	US-10-684-532-4	Sequence 4, Appli
43	662	19.1	751	US-09-795-693-14	Sequence 14, Appl
44	662	19.1	751	US-10-156-239-14	Sequence 14, Appl
45	662	19.1	751	US-10-199-485-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-10-437-963-140667
; Sequence 140667, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 140667
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41845C.1.pep
US-10-437-963-140667

Query Match 67.1%; Score 2330.5; DB 16; Length 657;
Best Local Similarity 70.6%; Pred. No. 8.4e-193;
Matches 441; Conservative 83; Mismatches 100; Indels 1; Gaps 1;
QY 43 QVEVPPQPFKSLKYSLETFFPDPPDLPQF-KNKPKSKFMLGLQFFFPFIPWPKYTF 101
DB 18 RVMPAKPFLETGLGNMKETFLDPPFRVRRRCGRRRAAALRYVFFPFWAHSYTL 77

Qy	102	QFLKADLIAGITIASLAIPOGISYAKLANLPILGLYSSFTPLPIYAMGSSRD LAVGTV	161
Db	78	GTLSKDLIAGITIASLAIPOGISYAKLANLPVLGLYSSFVPLVYAMGSSRD LAVGTV	137
Qy	162	AVCSLLMGWMLSNVDPNEDPKLYLHLAFTATLFAGVFOAALGHFRGLIVDFLSHATII	221
Db	138	AVASLLIGSMJSEEVSAADPALYLFHVALTATFFAGVFQALLGLRGLFIVDFLSHATIV	197
Qy	222	GFMGGAATVVCLQLQKLSILGLEBHFTHGADIIISVMRSVFOTHEWRWESAVILGCVFIFFLL	281
Db	198	GFMGGAATVVCLQLQKGMFGLDHFATTATDLVSMSSVFSQTHLWRWESVWVGCGFIFFLL	257
Qy	282	STRYFSKRPFRFWVSAMAPLTSVILGSLLYVFFHAEKHGVEVIGELKKGINPPSUTNLV	341
Db	258	ITRFFSKRPFRFWVSAAAPLASVILGSLLYVFFHAEKHGVOVIGLYLKGKLNPPSATS LN	317
Qy	342	FVSPYMTTAVKGTIVVGIISLAEGIAVGRSFAMYQNYINDGNKEMIAIGTMNVGSGFTSC	401
Db	318	FSSPYMMLAKTGITGVIALAEGIAVGRSFAMPKNYHIDGNKEMIAFGTMNVGSLTSC	377
Qy	402	YLTGTGFSRSANNYAGCKTAASNIIMSLAVMLTLFLTPLFHYTPVLVLSAIVGSAMLG	461
Db	378	YLTGTGFSRSANNYAGCKTAMSNVIMSVMYMTLLFLTPLFHYTPVLVLSAIVGSAMLG	437
Qy	462	LIDYEAAILHFKVDKDFVVCMSAYIGVVGSGVEIGLVAIVISVTLVLAFIARPTFVL	521
Db	438	LIDYPAAVHLWQDKVDFCVCLGAYLGVVGSGVEIGLVAIVGILVLAFLVAPRPTVL	497
Qy	522	GNIPNSVIYRNVEHYONAKHVPGLMILRIDAPIYFANASYLRERTWIDDEBERIKATG	581
Db	498	GNIPNTMIVYRMDOYTAQRVPGVLVLVVDSPVFTWASVYLRERIARWIDDEDOCKEKG	557
Qy	582	ETSLOQYIIDMGA VGNIDTSGISMJLEEYKKITERRELQVLVLPNVPSEVMKKLKNKSPONH	641
Db	558	EMGIQYVVLDMGAVGSDTSGTSMJDELRLKTLDRRGQLVLANPGSEIMKKLJSSKVLEA	617
Qy	642	LKKWYIYLTVEBAGVACNFURASK	666
Db	618	IGHWTIFPTVGEAVAECDPVMHSQK	642

```

RESULT 2
US-10-424-599-273958
; Sequence 273958, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 39-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273958
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89404C.1.pep
US-10-424-599-273958

```

	Query Match	58.0%	Score 2015;	DB 15;	Length 414;
	Best Local Similarity	94.5%;	Pred. No. 9.7e-166;		
	Matches 397;	Conservative	8;	Mismatches 9;	Indels 6; Gaps 1;
Qy	261	QTHEWRWESAVLGCVFIFLLSTRYFSKQKPRFFWVSAMAPLTSVILGSLLVYVTHAEKH	320		
Db	1	QTHEWRCESAVLGCCTFFLLVTRYFSKQKPRFFWVSAMAPLTSVILGSLLVYVTHAEKH	60		

QY	321	QVEVIGELKGLNPPSLNLNVFVSPYMTAVKTIWVGIIISLAEGIAVGRSFAMVKYNYI	380
DB	61	GVQ:-----KGLNPPSLNLNVFVSPYMTAVKTIWVGIIISLAEGIAVGRSFAMVKYNYI	114
QY	381	DGNKEMIAIGTMNVVGSFSTCYLITGPPSRSAVNNAGCKTAASNIIMSLAVMLTLLFLT	440
DB	115	DGNKEMIAIGTMNVVGSFSTCYLITGPPSRSAVNNAGCKTAASNIIMSLAVMLTLLFLT	174
QY	441	PLFHYTPLVLVSATIVSAMLGLIDYEAAIHLFKVDKDFVVCMSAYIGVVGSGVEIGLVI	500
DB	175	PLFHYTPLVLVSATIVSAMLGLIDYEAAIHLFKVDKDFVVCMSAYIGVVGSGVEIGLVI	234
QY	501	AIIVSLVRLLFIARPTFVLGNIPNSVIYRNVHEYQNAKHVPGMLILEIDAPIYFANAS	560
DB	235	AIIVSLVRLLFIARPTFVLGNIPNSVIYRNVHEYQNAKHVPGMLILEIDAPIYFANAS	294
QY	561	YLREIRITRWIDEEBERIKATGETSLOQYVIIDMSAVGNITDTSISMLEEVKKITERRELQL	620
DB	295	YLREIRITRWIDEEBERIKATGETSLOQYVIIDMSAVGNITDTSISMLEEVKKITERRELQL	354
QY	621	VLVNPVSEVMKKLNSKFQNLGKKWIYLTVEEAVGACNENLRASKTNPKKDETEGNNV	680
DB	355	VLVNPVSEVMKKLNSKFQNLGKKWIYLTVEEAVGACNENLRASKTNPKKDETEGNNV	414
RESULT 3			
US-10-425-115-235264			
; Sequence 235264, Application US/10425115			
; Publication No. US20040214272A1			
; GENERAL INFORMATION:			
; APPLICANT: La Rosa, Thomas J.			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants			
; FILE REFERENCE: 38-21(53222)B			
; CURRENT APPLICATION NUMBER: US/10/425.115			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 369326			
; SEQ ID NO 235264			
; LENGTH: 695			
; TYPE: ERT			
; ORGANISM: Zea mays			
; FEATURE:			
; OTHER INFORMATION: Clone ID: MFT4577_146143C.1.pep			
US-10-425-115-235264			

	Query Match	53.3%;	Score 1853.5;	DB 16;	Length 695;
	Best Local Similarity	56.3%;	Pred. No. 2e-151;		
	Matches	346;	Conservative 115;	Mismatches 153;	Indels 1; Gaps 1;
Qy	42	HQVEVPPQPPFSLKYSLKSETPPDPLRQPKNKPKASKKFMGLGLOFFFPFIWEWAPKYPF	101		
Db	70	YKVGVPPEKNLLAEISDAVKETFPADPLRQYKDQPSKKIWLGLQHFIPVLEWSRHSYL	129		
Qy	102	QFLKADLIAGITITASLAIPOGISYAKLANLPPIGLLYSSFPPIIYAMMGSSRDLAIVGTV	161		
Db	130	GKFKGDIAGITITASLCIPDQIGYSKLANUPAEVGLYSSFPVPLIYAMMGSSRDIALGPV	189		
Qy	162	AVGSLLMGSMLSNAVDNEDPKLYHLHAFATATLFAGVFQAALGFLRLGLIVDFLSHATII	221		
Db	190	AVVSLLIAGITLQNEIDPKTHLEPYRLAFTATFPAGVTQAALGFFRLGFIIEFLSHAAIV	249		
Qy	222	GFWGGAATVCLQOLSKSILGLEHFTHGADIISWRSVFTOTHE- WRWESAVLGCVFIFPL	280		
Db	250	GFWAGAAITIALQOLKGLFIANTFTKKSIDIVSYMKSVGWNVHHGWNWQTILIGATFLAFL	309		
Qy	281	LSITRYFSKKEPRPFWSAMAPLNSVILGSLLVFTTHAEKHGVEVIGELKKGLNPPSITNL	340		
Db	310	LVAKYIGKRNKKLUPFWSAIAPLTSVITSTFVYITRADHGVAIVKRNKINGINPPPSASLI	369		
Qy	341	VFVSPYMTTAVKTGIVVGGIISLAEGIAVGRFSFAMKYKNYNIDGNKEMIAIGTMNVGSGPTS	400		

```
Db 370 YFTGPIATGFKGIVAGMIGLTAIAIGRTFAALKDYRIDGNKEMMALGTMINIVGSLTS 429
Qy 401 CYLTGTPFSRSVAVNAGCKTAASNIIMSLAVMLTLFLPLFHYTPLVVLIAIIVSAML 460
Db 430 CYVATGSFSRSVAVNAGCKTAVSNVMSVVMTLTLITLPLFKYTFENAILSSIIISAVL 489
Qy 461 GLIDYEAIAHLFKVDKDFVVCMSAYIGVVGSGVEIGLVIATVIVSVLRLFTIARPTFV 520
Db 490 GLIDYESAYIWKVDKDLFLACMGAFGFFSSVEYGLLIAVLSAKILLQVTRPTVL 549
Qy 521 LGNIPNSVIYRNVEHYQNAKHVPGMLTLEIDAPIYFANASYLRIRTRWIDEEERIKAT 580
Db 550 LGNLPRTIYRNVEYQPDATKVPGLIVRDSAIYFTNSVYKVERILRWLRDSEEQDDQ 609
Qy 581 GETSLQVVIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNSKFN 640
Db 610 KLTKTEPLIVDLSPVIDIDTSGIHAELEAKLEKRIQLVLTNPGPAVIQKLSAKFTD 669
Qy 641 HLGKKWIYLTVEAV 655
Db 670 MIGEDNIFLTVGDAV 684

RESULT 4
US-10-424-599-207945
; Sequence 207945, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 207945
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(659)
; OTHER INFORMATION: unsure at all xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29800C.1.pep
US-10-424-599-207945

Query Match 53.2%; Score 1850; DB 15; Length 659;
Best Local Similarity 55.9%; Pred. No. 3.8e-151;
Matches 359; Conservative 122; Mismatches 153; Indels 8; Gaps 4;

Qy 41 VHQVEVPPPPFFKSLKYSKLTFFPDPLRQPKNKPKASKFMLGQFFPPPIFEWAPKYT 100
Db 20 VHQVPPPHKSTLQKLKRLKETFFPDPLRQPKGQPLKRLILGAYVPPILQWGPKN 79
Qy 101 FQFLKADLIAGITIASLAIPQGISYAKLANLPITGLYSSFIPIPLIYAMGSSRDIAVGT 160
Db 80 LKLFKSLVSLGLTASLDIPQGSYAKLASLPILGLYSSFVPPVAVLNGSSKDLAVGP 139
Qy 161 VAVGSLMGSLMSNAVDNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLGHATI 220
Db 140 VSTASLVMSGLMQHVEVSPTTDPILFLQLAFTSLTFLAGLFOALLGILGLFIIDFLSKAIL 199
Qy 221 IGFMGGAATVVCVQLKLSILGLEHFTHGADIISVMRSVFTQTHWRWESAVLCGVFIPL 280
Db 200 IGFMGAAATVLSQLKSLILGITHFTNQMGILIPMTSVFNHIEHWSQOTILMGICPLVIL 259
Qy 281 LSTRYSKKEPRFVWSAMAPLTSVILGSLLVFTHAEKHGVEVIGELKKGLNPPSLNL 340
```

```
Db 260 LLAHVSIIRKPKLFWVSAGAPLMCVIISTLLVFAIKAQNHGISAIKLGQOINPPSPNNML 319
Qy 341 VFVSPYMTAVTKTIGIVVGIISL-AEGIAVGRSPAMTKYNNIDGNKEMIAIGTNNVVGSGFT 399
Db 320 LPHGSHLGLVMKTGLTIGILSTXEGIAVGRTEFAALKKNYKVDGKNEMMAIGFNNVVGSGFT 379
Qy 400 SCVLTTPSPRSVAVNAGCKTAASNIIMSLAVMLTLFLPLFHYTPLVVLIAIIVSAML 459
Db 380 SCVYTTGAFSRSVAVNAGCKTAVSNVMSVVMTLFLPLFQYTPNVVLGAIIIVTAV 439
Qy 460 LGLIDYEAIAHLFKVDKDFVVCMSAYIGVVGSGVEIGLVIATVIVSVLRLFTIARPTF 519
Db 440 IGLIDLPACANIWKIDKDFVVMWMTAFGLVLFISVGGGLALAVGLSTLKLILQITRPTKV 499
Qy 520 VGNIPNSVIYRNVEHYQNAKHVPGMLTLEIDAPIYFANASYLRIRTRWIDEEERIKAT 579
Db 500 MLGKIPGTDIYRNLDQYKEAVRIPGFLIISIEAPINFANITYLNERTLRWIDEEEDNIKE 559
Qy 580 TGSETSLOVVIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNSKFK 639
Db 560 --QLSLRFLVLEMSAVSVDTSGLSLFKELKATLEKKGVLELVNPLAEVIEKLKKADEA 617
Qy 640 N-HLGKKWIYLTVEAVGACNPNLRASKTNPKKDETEGWNNV 680
Db 618 NDFIRADNLFITVGEAVA-----SLSSAMKGQSSTITEGAHTI 655

RESULT 5
US-10-437-963-198336
; Sequence 198336, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198336
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94005C.1.pep
US-10-437-963-198336

Query Match 52.1%; Score 1811.5; DB 16; Length 666;
Best Local Similarity 53.8%; Pred. No. 8.3e-148;
Matches 340; Conservative 122; Mismatches 167; Indels 3; Gaps 3;

Qy 43 QVEVPPPPFFKSLKYSKLTFFPDPLRQPKNKPKASKFMLGQFFPIFEWAPKYT 102
Db 36 EVNLGRRFPFAELKWSLDLAETFFPDPPFGFALPPARRAWCAKTFVPAIDWVPYGLD 95
Qy 103 FLKADLIAGITIASLAIPQGISYAKLANLPITGLYSSFIPIPLIYAMGSSRDIAVGTVA 162
Db 96 KFKFDLLAGITIASLAIPQGISYARLANLPITGLYSSFPVPLMYAVFGSSNNLAVGTVA 155
Qy 163 VGSLLMGSLMSNAVDNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLGHATIIG 222
Db 156 AASLLIASIETEAADENPOLYQLFYTAAPFTGLFQALGVFRGLIVDFLSRSTITG 215
Qy 223 FMCGAATVVCVQLKLSILGLEHFTHGADIISVMRSVFTQTHWRWESAVLCGVFIPL 282
Db 216 FMGGTAMIIILQOFGKLLGKHFHTTKTDIIISVLHSTYHWRHKEWQSAVLGICFLFLMS 275
```


[illegible]

Db	155	VS	AS	LV	MG	SM	L	R	Q	A	V	S	P	D	E	P	I	L	Y	L	Q	L	A	F	T	S	T	F	F	A	G	V	F	Q	A	S	L	G	F	L	R	L	G	I	P	V	D	F	L	S	K	A	T	L	214				
Qy	221	I	G	F	M	G	G	A	A	T	V	V	C	L	Q	L	K	S	I	L	G	L	E	H	F	T	H	G	A	D	I	I	S	M	B	S	V	F	T	Q	T	H	E	R	W	E	S	A	V	L	G	C	F	I	F	F	L	280	
Db	215	T	G	F	M	G	G	A	A	I	V	S	L	Q	L	K	G	L	L	I	I	H	F	T	S	Q	M	G	F	V	Q	M	S	V	F	K	H	D	E	N	A	Q	T	I	L	M	G	V	A	F	L	A	V	L	274				
Qy	281	L	S	T	R	S	F	K	K	R	P	F	F	W	S	A	M	A	P	L	T	S	V	I	L	G	S	L	L	V	T	H	A	E	K	H	G	V	E	I	G	E	L	K	K	L	N	P	P	S	L	T	N	L	340				
Db	275	L	T	R	H	I	S	A	R	N	P	K	L	F	W	S	A	A	A	P	L	T	S	I	T	S	I	F	S	V	K	A	--	H	G	I	S	V	I	G	D	L	P	K	L	N	P	P	S	A	N	M	L	332					
Qy	341	V	F	S	P	M	T	T	A	V	K	I	V	V	G	I	I	S	L	A	E	G	I	A	G	B	S	F	A	M	Y	K	N	I	D	G	N	K	E	M	I	A	I	G	T	M	V	V	S	F	T	S	400						
Db	333	T	F	S	G	S	V	G	L	A	N	T	G	I	M	T	G	I	L	S	L	T	E	G	I	A	G	T	F	A	S	I	N	N	Y	Q	V	D	G	N	K	E	M	A	I	G	V	M	N	A	G	S	C	A	S	392			
Qy	401	C	Y	L	T	T	G	P	F	S	R	S	A	V	N	Y	N	A	G	C	K	T	A	A	S	N	I	M	S	L	A	V	M	L	T	L	F	L	T	P	L	P	H	Y	T	P	L	V	L	S	A	I	I	V	S	A	M	L	460
Db	393	C	Y	T	T	G	S	F	S	R	S	A	V	N	Y	N	S	A	G	C	T	A	S	N	I	M	S	A	V	L	T	L	F	L	M	P	L	P	H	Y	T	P	N	I	L	S	A	I	I	T	A	V	I	452					
Qy	461	G	L	I	D	Y	E	A	A	H	L	F	K	D	K	P	F	V	C	M	S	A	I	G	V	V	G	S	E	I	G	L	V	I	A	I	V	I	S	V	L	R	V	L	L	F	T	A	R	P	R	T	F	V	520				
Db	453	G	L	I	D	V	R	G	A	A	R	L	K	V	D	L	F	L	A	C	M	A	A	F	L	G	V	L	L	S	V	Q	M	G	L	A	I	A	V	G	I	S	L	F	K	I	L	L	Q	T	R	P	N	M	V	512			
Qy	521	L	G	N	I	P	S	V	I	R	N	E	H	Q	N	A	K	H	V	G	M	L	I	E	D	A	I	P	I	F	A	N	A	S	A	L	E	R	I	T	R	M	I	D	E	E	E	R	I	K	A	T	580						
Db	513	K	G	V	P	G	T	S	R	S	Y	S	Q	Y	R	E	A	R	M	R	V	P	S	F	L	V	G	V	E	S	A	I	Y	F	A	N	S	M	Y	L	G	E	R	I	M	R	F	L	E	E	D	E	R	A	A	K	572		
Qy	581	G	E	T	S	L	Q	V	I	D	M	S	A	V	G	N	I	D	T	S	G	I	S	M	L	E	E	V	K	I	T	E	R	R	E	L	Q	L	V	L																			

Db 155 VSISLIMGAPCCVAGREPHRGADAVPAAGLH-----XTLFAGLVQASLGLRLRGFVIDFL 210
Qy 216 SHATIIIFGMAATVVCLOQLKSLGLEHETHGADIISVNRSVFTQTHWRWESAVILGCV 275
Db 211 SKATLVGFMAAGAIIVALKQKALLGIVHFTTEMGIVPVMAVVFHHTSEWSQTIILMGVC 270
Qy 276 FIFELLSTRFYSKRPFFVWSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGILNPP 335
Db 271 FLVFLLSARHVSIRWPKLFWVSACAPLASVTISTLLVFLFKAQNHGSIIGOLKCGLNRP 330
Qy 336 SLTNLVPVSPMTTAVKTIWVGIIISLAEGIAVGRSPAMKYNIDGNKEMIAIGTMNVV 395
Db 331 SWDKLLFDYALGUTKTIWVGIIISLAEGIAVGRSPAMKYNIDGNKEMIAIGTMNVV 390
Qy 396 GSFTSCYLTTGPPRSRNVNAGCKTAASNIIMSLAVMLTLLFPLFHYTTPVLVLSAII 455
Db 391 GSCTSCVTTGAFSRSVAVNAGCKTAMSNVIMALTVMVTLFLMPLFVYTPNVVLGAI 450
Qy 456 VSAMGLIDYEAIAHLFKVDKDFVVCMSAYIGVWFGSVEIGLVAIVISVLRLFIAR 515
Db 451 IAAVIGLIDFPVAVHVKMDKDFVVCMSAYIGVWFGSVEIGLVAIVISVLRLFIAR 510
Qy 516 PRFVLGNIPNSVIRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRERITWIDEEBE 575
Db 511 PKMVQNIIGTDIYRDLHYKQVRVSGFLIILAEPINFANSYLNRIKRWI--EEE 568
Qy 576 RIKATGETSLQYIIDSAGVNIIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLKN 635
Db 569 SFEQDKHTLHFIILDSAVPAIDTSGIAFLIDIKKSIERRGLELVLPVNPTEGXHGNTT 628
Qy 636 SK--FQNLGKKMIVLTVEEAVGACNPNLRASKTNP 669
Db 629 CKRGKXITLGGIACIXTTGEAI--ASLSALAKMTKP 662

RESULT 12

US-10-437-963-174545
; Sequence 174545, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)/B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174545
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72475C.1.pap
US-10-437-963-174545

Query Match 45.0%; Score 1563.5; DB 16; Length 659;
Best Local Similarity 47.7%; Pred. No. 2.6e-126;
Matches 299; Conservative 121; Mismatches 194; Indels 13; Gaps 3;
Qy 50 QPFPKSLKYSKLTFFPDPLRQFNKPKASKKFMGLQFPFFIFEWAPKTYTFOFLKADLI 109
Db 29 RPFCAARAALKDITLFPDPPFRGLGCMPPARRAWRVFVPALDWGAGSAASFWYDL 88
Qy 110 AGTITIASLAIPOGISYAKLANLPILGLYSFPLIYAMGSSRLAVCTVAVGSLMG 169

Db 89 AGVTIASLAIPOGISYATLAGIPPVILGYSCFVPPVLYAVMGSSRNLGVPVATSSLLVA 148
Qy 170 SMLSNAVDNEDPKLHLHLAFTATLTFAGVQFQAALGLFRLGLIYDFLSHATIIIFGMGAAT 229
Db 149 SIUGGKVRASDDQRLYLTQVFTSAPFTGLVQALGLRLGLIYDFMSPRPAITFGMGTAI 208
Qy 230 VVCLQOLKSLGLEHETHGADIISVNRSVFTQTHWRWESAVILGCVFIFELLSTRFYSK 289
Db 209 VIMLOQLKFLGTHFTTKTIDIVSLRYIFHNTHOMQOSTVLGVCFILFVFTQVRRR 268
Qy 290 RPRFFVWSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGILNPPSLTNLVPVSPYMTT 349
Db 269 RPKLFWVSAMSPLLVVVGVCFSLIKKHGPIVGTLKRGINPSSISOLKEQPEVGV 328
Qy 350 AVKTIGVIGIISLAEGIAVGRSPAMKYNIDGNKEMIAIGTMNVVSGFTSCYLTTGPPS 409
Db 329 AMKAGFVSGMLALAEVAVGRSFAAMKERRIDGNKEMVAFGLMNLIGSFTSCYITGAPS 388
Qy 410 RSNVYNAGCKTAASNIIMSLAVMLTLLFPLFHYTTPVLVLSAIIIVSAMGLIDYEA 469
Db 389 KTAVNYHAGCRTAMSNVMSVCMALVLALPLFRHTPLVALAAIITSSMLGLVKGREIR 448
Qy 470 HLPKVDKDFVVCMSAYIGVWFGSVEIGLVAIVISVLRLFIARPRFTFVLGNI----- 524
Db 449 RLYEVDKADFVCAALGLVVFSTMITGLGVAIVISVLRLHVARPSTKLGVRVCGSG 508
Qy 525 ----PNSVIRNVEHYQNAKHVPGMLILEI-DAPIYFANASYLRERITWIDEEBEIRKA 579
Db 509 AGAADDHAFCDVAQPGAATAPSLVLQVAGSPVCFANAEYLRLERIRARWDEE---KA 565
Qy 580 TGSETSLOQYIIDSAGVNIIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLKNK 639
Db 566 VAGEDLLYVLDIGGVTAIDSPGIELRVEHGBLERKGMKMAVTNPRMAVAKLVLGSLA 625
Qy 640 NHLGKKMIVLTVEEAVGACNPNLRASK 666
Db 626 ELVGESWFMFLSNGDALLAACRYTLQSGK 652

RESULT 13

US-10-424-599-195017
; Sequence 195017, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)/B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 195017
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_18127C.1.pap
US-10-424-599-195017

Query Match 43.2%; Score 1500; DB 15; Length 653;
Best Local Similarity 48.0%; Pred. No. 8.1e-121;
Matches 297; Conservative 126; Mismatches 178; Indels 18; Gaps 7;
Qy 48 QPFPKSLKYSKLTFFPDPLRQFNK-----PASKKFMGLQFPFFIFEWAPKTYT 100
Db 38 PPSF-WRVVADSVSKT-----ISHYKHLSSLIDQPCCTLLSLVQLVFPVPLAWGRNTY 90
Qy 101 FQPLKADLITAGTITIASLAIPOGISYAKLANLPILGLYSFPLIYAMGSSRLAVCT 160
Db 91 ATKFRKOLLAGLTIASLCIPQSIGYATLAHLDPQYGLTYSVPPPLIYAVMGSSREIAIGP 150

```

161 VAVGSLMGSMLSNAVDPNEDPKLVLHLAFTATLTFAGVFOAALGALFRIGLIVDFLSHATJ 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 VAVVSLLSMMKEKLVDPATDPVGVYTKLILLATLTFAGIFQTSFGLRLGLFLVDLFLSHAAI 210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
221 IGFWGGAAVTVCLQQLKSIILGLEHFTGADIIISVMRSVPTQTHE-WRWESAVLGCVPFF 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
211 VGFVAGAAIVLGLQQLKGLLGTHTFTKTDIVSVWKAVMVEAHNPWNPNRNFILGCSFLVF 270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
280 LLSTRYFSKRPFRFFWWSAMAPLTSVILGSLVIVYTHAEKHGVEVIGELKKGLNPPSLTN 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
271 ILTRFLGKRKKLFWLASISPLVSVLSTLIVELTRADKNGVKIVKHVGKGLNPSLLHQ 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
340 LVFVSPYMTTAVKTIIVGIIISLARGIAVGRSFAMKYNNYIDGNKEMIAIGTMNVVGSPT 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
331 LDFNPNYIGEYAKIGLVVAVVALTESIAGVRSFASIKGYQLDGNKEMSIGLNTIIGSFT 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
400 SCYLLTGPFSRNVNNAAGCKTAASNITMSLAVMLTLPLTLPHYTLPLVLSAIIIVSAM 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
391 SCYVATGSFSAIVNVAAGCETVLVNNIVMAITVLISLQFLTLLYTPAILASVILSAL 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
460 LGLIDYEAHILFKVDKDFVVCMSAYIGVWFGSVYEIGLIVAIIVSLVLLFIARPTFF 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
451 PGLIDLSEAYKIMWKIDIDFLACAGAFGVLFPASVEIGLVAVVIFSFKIILISIRPGTE 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
520 VLGNIPNSVIVRNVEHYQNAKHVPGMLLEI-DAPIYFANASYLBERITRWIDESE-ERI 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
511 TLGKIPGDLFCDVQYVPMAVKVPGVMILIRVKSALLCFANANFVERIRIKWYTESEDD 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
578 KATGETSLQYIIVDMSAVGNITDTSIGSMLEVEKKLITERRELQVLVNPVSEYMKKLNKSK 637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
571 KGNRSRTIQLVILDTSLNVLNIDTSGITALEELHKSLSQGKQLATANPFWQVIHKLKVSN 630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
638 FQNHGKKWIVYLTVEEAVG 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
631 FVGKIGGR-VELTVEEAVG 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 14
US-10-425-114-69272
; Sequence 69272, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21153313/B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69272
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMST02400016G05_FLI.pep
US-10-425-114-69272

```

	Query Match	42.0%	Score 1460;	DB 15;	Length 575;
	Best Local Similarity	49.3%;	Pred. No. 2e-117;		
	Matches 281;	Conservative 120;	Mismatches 165;	Indels 4;	Gaps 4;
Qy	90	FPFEWAPKYTEFQLKADLTIAGITIASIATPQGISYAKLANLPILGLYSSFIPLIYAV	149		
		: :			
Db	2	FPIAWGNVTATFKRDKLLAGLIASLCIPQSGYATLAHLDROYGLIYSVPLIYAV	61		
		: :			
Qy	150	MGSSRDIAVGTVAVGSLMLMSLSNAVDPNEDPKLYHLHAFTATLFAGVFQAALGLFLRG	209		
		: :			

```

Db      62  MGTSRETAIGPVAVVUSLLSSMMEKLVDPATDPGVGYTKLIIATLTFAGIFQTSFGGLRLG  121
Qy      210  LIVDFLSHATIIIGFMGGAATVVCLOQLKSIIGLEHFTHGADIISVMRSVFTQTHE-WRWE  268
Db      122  FLVDFLSHAAIVGFVAGAAIVIGLQQLKGLLGIHTFTTKTDIVSVKMAWEAVHNPMPNR  181
Qy      269  SAVLGCYFIPELLSTRYFSKKRPFVWSAMAPLTSVLGSLIYYPTHAEKHGVEVIGEL  328
Db      182  NFILGCSFLVFILTRCLGRKCKKLFWLASISPLSVSVVSTLIIVFETRADKNGKXIVXHV  241
Qy      329  KKGLENPSELNLVFPSPYMTTAVKTGIWVGIISIAELGAVGRSFAMYKNYKYNIDGNKEMIA  388
Db      242  KGLENPSSIHOLEDNNPYIGEVAKIGLVAVVAVLTESAVGRSPASIKGYQLDGNKEMMS  301
Qy      389  IGTWNVVGSFTSCYLTTPGFSRSVANNAGCKTAASNIIMSLAVMLTLFLTPPLPHYTPL  448
Db      302  IGLTNIIGSFTSCYVATGSPFRTAVNYAAGCETLVSNIVMAITVLISLQFTKLIIYYTPT  361
Qy      449  VVLSAIIIVSAMGLIDYEAATHLFPKDKDFVCMGSAVIGVVGFSVEIGLVIATIVSVLR  508
Db      362  AILASVILSAUPLGILIDSEAYKIWKVDKIDFLACAGAFGLVPASVEIGLVAVVVISFSK  421
Qy      509  VLLFIARPTPLGNIPNSVIYRNVEHYQNAKHVPGMLILBI-DAPIYFANASYLRERIT  567
Db      422  IILSIRPGTETLCKLPGTDLFCDVQYQYPMAVKVPGVMIIRVKSALLCFANANFVRERII  481
Qy      568  RWIDEEB-ERIKATGETSLOVVIDMSAVGNIDTSGISMLSEVKKITERRELOLVVNPV  626
Db      482  KWTVEEEDDDKGNRSRTIQFLIILDTSNLVNIDTAGITALELHKLSLSHGKQLAIANPR  541
Qy      627  SEVMKLNKSKFQNHGKKGKIYLTVEEAVG  656
Db      542  WQVTHKLKVSFNVCVKIRGR-VFLIVEEAVG  570

RESULT 15
US-10-425-115-228851
; Sequence 228851, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 228851
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_140304C.1.pep
US-10-425-115-228851

```

Query Match	41.3%;	Score 1436;	DB 16;	Length 405;
Best Local Similarity	67.3%;	Pred. No. 1.5e-115;		
Matches 268;	Conservative 65;	Mismatches 61;	Indels 4;	Gaps 1;

Qy	275	VFIIFLLSTRYFSKRRPRFFFWSAMAPLTSVILGSLLYVFTFAEKGVVEVIGELKKGLNP	334
Db	3	IICFFFIQ----SKRPKPLFWISAAAPLTSVVLGSLVYLTHAENHGVIEVIGLYLKKGLNP	58
Qy	335	PSLNLNVFVSPYMTTAVKTGTIVGIIISIAEGIAVGRSPAMYKNYINDGNKEMIAIGTMNV	394
Db	59	PSVTSLOFPPYMMIAIKGTGIITGVIALAEGIAVGRSPAMFNHYIDGNKEMIAIGTMNV	118
Qy	395	VGSFTSCVLTITGPFSSRAVNNAGCKTAASNIIMSLAVMLTLLFUTPLFPHYPLVLSAI	454
Db	119	LGSITSCVLTITGPFSSRAVNNACRTAMSNVMSLAVMLTLLFUTPLFPHYPLVLSAI	178

Qy	455	IVSAMLGLIDYERAIHLPKYDKEDFVVCMSAYIGVWFGSVBIGLVIAIVISVLRVLLFIA	514
Db	179	IVSAMLGLVDFGAALHLWRVDKDFVCAGAYLGWFGSVVEGLWAVASLLRVRFGA	238
Qy	515	RPRTFVLGNIPNSVIYRNVEHYONAKHVPGBMLILEIDAPIYFANASYLRERITRWIDEE	574
Db	239	RPRTVLGNIPGTMVYRMDQYAAAQTVPGVLVRVDAPVYFANASYLRERISRWIDEE	298
Qy	575	ERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQLVLVNPVSEVMKKLN	634
Db	299	ERTKSQCEMGVRYVWLDMGAIGSIDTSGTSMLELNKSLDRRGWQIVLANPGSEIMKKLD	358
Qy	635	KSKFQNLGKKWIYLYTVEEAVGACNPNLRASKTNPKD	672
Db	359	SSKVLQIGHWVPPTVGEAVASCDYVLHSHKPGMAKD	396

Search completed: August 31, 2005, 11:53:40
Job time : 59 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2005, 08:46:03 ; Search time 35 Seconds
(without alignments)
1869.354 Million cell updates/sec

Title: US-10-762-049-18
Perfect score: 3475
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKKDETEGWNV 680

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2642.5	76.0	658	2 T48901	sulfate transport
2	2641.5	76.0	658	2 T49069	sulfate transport
3	2638.5	75.9	703	2 T51161	hypothetical prote
4	2575.5	74.1	646	2 T48902	sulfate transport
5	2314	66.6	646	2 T01079	sulfate transport
6	1898.5	54.6	631	2 B86365	probable sulphate
7	1796	51.7	683	2 B86354	protein F282.22 [i
8	1758.5	50.6	719	2 D96809	protein F28K19.22
9	1752	50.4	649	2 T00946	probable sulfate t
10	1751.5	50.4	649	2 T51839	sulfate transport
11	1749.5	50.3	660	2 T04416	sulfate transport
12	1735.5	49.9	662	2 S51766	sulfate transport
13	1722.5	49.6	667	2 S51763	sulfate transport
14	1507	43.4	677	2 T50022	sulfate transport
15	1503	43.3	644	2 S51765	sulfate transport
16	1392.5	40.1	658	2 S74246	sulfate transport
17	998	28.7	485	2 S34800	sulfate transport
18	840.5	24.2	750	2 B90137	sulfate permease [
19	690.5	19.9	573	2 D83440	probable sulfate t
20	687.5	19.8	233	2 T01205	sulfate transport
21	658	18.9	566	2 S74633	high affinity sulf
22	658	18.9	605	2 G70409	high affinity sulf
23	658	18.9	840	2 T39116	probable sulfate p
24	649.5	18.7	877	2 T40413	sulfate permease -
25	637.5	18.3	739	2 A54808	diastrophic dyspla
26	618.5	17.8	553	2 AE1140	probable sulfate t
27	615	17.7	560	2 F70688	hypothetical prote
28	613.5	17.7	553	2 A11498	probable sulfate t
29	598	17.2	749	2 T23628	hypothetical prote

30	586.5	16.9	764	2 A47456	down-regulated in
31	578	16.6	893	2 S64926	probable membrane
32	571.5	16.4	703	2 A49994	sulfate transport
33	565.5	16.3	611	2 T27820	hypothetical prote
34	563	16.2	754	2 S52816	probable membrane
35	562	16.2	592	2 B82498	sulfate permease f
36	559.5	16.1	859	2 S46176	sulfate transport
37	554.5	16.0	700	2 T23629	hypothetical prote
38	510	14.7	650	2 T16077	hypothetical prote
39	481.5	13.9	667	2 T41306	probable sulfate t
40	467	13.4	544	2 AH1544	transport protein
41	463.5	13.3	758	2 T25751	hypothetical prote
42	451	13.0	537	2 F97546	hypothetical prote
43	451	13.0	537	2 AB2766	hypothetical prote
44	448	12.9	517	2 C83642	sulfate permease [
45	430	12.4	541	2 A11186	probable sulfate t
					transport proteins

ALIGNMENTS

RESULT 1

T48901
sulfate transporter ATST1 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cross)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T48901
R;Sohlberg, L.E.; Sussex, I.M.
Plant Physiol. 113, 1463, 1997
A;Title: Nucleotide sequence of a cDNA (Accession No. U52970) encoding a Cys proteinase f
A;Reference number: Z15626
A;Accession: T48901
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-658 <SOH>
A;Cross-references: UNIPROT:Q9SVL3; EMBL:D89631; PIDN:BAA21657.1
A;Experimental source: Landsberg erecta; tissue type above-ground parts
C;Genetics:
A;Gene: ATST1
C;Superfamily: sulfate transport protein

Query Match	76.0%;	Score 2642.5;	DB 2;	Length 658;
Best Local Similarity	76.0%;	Pred. No. 2e-174;	Mismatches 80;	Indels 3;
Matches 501;	Conservative	75;	Gaps 3;	
QY	24	MGSDVYEYPLGMNFERVHQ-VEVPPQPFKSLKYSKLETFFPDPLRQFNKPASKKF	82	
DB	1	MGTEDTYFPQGAELHRRHHTVEAPQPQFLKSLQYSVKETLFPDDPFRQFNQNASKRF	60	
QY	83	MLGLQFPFPIFEWAPKYTFQFLKADLIAGITIASLAIPOGISYAKLANLPPILGLYSSPI	142	
DB	61	VLGLKYFLPIFEWAPRYNLKFFKSDLIAGITIASLAIPOGISYAKLANLPPILGLYSSFV	120	
QY	143	PPLIYAMGSSRDALVGTAVGSLMGSLSNADVDPKLYLHLAFTATLTFAGVFOAA	202	
DB	121	PPLIYAVLGSSRDALVGTAVASLLTGAMLSKEVDAEKOPKLYLHLAFTATLTFAGVLEAS	180	
QY	203	LGFLRLGLIVDFLSHATIIIGFMGGAATVVCLOQLKLSILGLEHFTHGADIIISVMSRVSQT	262	
DB	181	LGIFRLGFIIVDFLSHATIVGFMGGAATVVSLOQLKGI FGLKHFTDSTDIVISVMSRVSQT	240	
QY	263	HEWRWSAVLGCVFIFFLIISTRYFKRPRFPFVWSAMPLTSVILGSLLYFYFTHAEKHGV	322	
DB	241	HEWRWSVGLGCGFLFLLSTRYFSIKKPKFFWVAAMPLTSVILGSLLYFYFTHAEKHGV	300	
QY	323	EVIGELKKGNLPPSLTNLVPVSPYMTAVKTGVVGIISLAEGIAVGRSPAMVKNYNIDG	382	
DB	301	QVIGDILKKGNLPLUGSDLIFTSPYMTAVKTGIIITGLIAEGIAVGRSPAMVKNYNIDG	360	
QY	383	NKEMIAIGTMNVVGSFVTSVLTGTFPSRAVNNAGCKTAASNIIMSLAVMLTLLFLTPL	442	
DB	361	NKEMIAFGMNVVGSFVTSVLTGTFPSRAVNNAGCKTAMSNIVMAIWMFTLLFLTPL	420	

```
QY 443 FHYTPLVLSAIIIVSAMGLIDYEAAILHLFKVDKDFVVCMSAYIGVFGSVVEIGLVIAI 502
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 FHYTPLVLSAIIISAMGLIDYQAAIHLWKVDKDFDLVCMSAYVGVFGSVVEIGLVIAV 480
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 503 VISLVRLLFIARPRFTVLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYL 562
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 AISIARLLLFVSRPKTAVKGNIPNSMIYRNTQYPSRTPVFGILLILEIDAPIYFANASYL 540
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 563 RERITWIDEEERIKATGETSLQYVIIDMSAVGNIDTSGISMLBEVKKITERRELQVL 622
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 RERIIRWIDEEERVKQSGESSLQYIILDMASVGNIDTSGISMMVEIKKVIDRRALKVL 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 623 VNPVSEVMKLNKSKF-QNHGLGKKWYLTVEEAVGACNFNLASKTNPKKDETEGNNV 680
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 601 SNPKGEVVKLTRSKFTGDHLGKEMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNNV 658

RESULT 2
T49069
sulfate transporter (AT5T1) - Arabidopsis thaliana
N:Alternate names: protein F4F15.10
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T49069
R:Alcaraz, J.P.; Clabault, G.; Cottet, A.; Mache, R.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25015
A:Accession: T49069
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-658 <AC>
A:Cross-references: UNIPROT:Q9SV13; EMBL:AL049711; GSPDB:GN00061; ATSP:F4F15.10
A:Experimental source: cultivar Columbia; BAC clone F4F15
C:Genetics:
A:Gene: ATSP.F4F15.10
A:Map position: 3
A:Introns: 116/1; 185/2; 242/3; 264/3; 302/3; 341/3; 384/1; 479/3; 543/2; 572/1; 597/3
C:Superfamily: sulfate transport protein

Query Match 76.0%; Score 2641.5; DB 2; Length 658;
Best Local Similarity 75.9%; Pred. No. 2.4e-174;
Matches 500; Conservative 81; Mismatches 75; Indels 3; Gaps 3;

QY 24 MGSVDYEYPLGMNFRVHQ-VEVPPPPFFKSLKYLKETFPPDPLRQFKNKPASKKF 82
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MGTEDTFFQGAELHRRHHTVEAPQPPFLKSLQYSVKETLFPDDPPFQFKQNQASRKF 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 83 MLGLQFFPIFEWAPKTYTQFLKADLIAGITIASLAIPOGISYAKLANLPPILGLYSST 142
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 VLGLKYFLPIFEWAPRYNLKFFKSDLIAGITIASLAIPOGISYAKLANLPPILGLYSST 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 143 PPLIYAMGSSRDLAGVTAVGSLMGSLMNAVDNEDPKLYHLAFTATLPAFVFOAA 202
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 PPLUYAVLGSSRDLAGVTAVASLLTGAMLSKEVDAEKDPKLYHLAFTATLPAFVLEAS 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 203 LGLFRGLIVDFLSHATIIIGFMGGAATVVCLOOLKSLILGLEHFTHGADIISVMRSVFTQT 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 LGIFRLGFIIVDFLSHATIVGFMGGAATVVSLOQLKGI FGLKHTDSTDVISMRSVFSQT 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 263 HEWRWESAVLGCVFIFLLSTRYFSKRPFRFFWVSAMAPLTSVILGSLLVYFTHAEKHGV 322
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 HEWRWESVGLGCCFLFLLSTRYFSIKKPKFFWVAAMAPLTSVILGSLLVYFTHAEKHGV 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 323 EVIGELKGLNPPSLNLVPSYMTTAVKTVGVGIISLAEGIAVGRSFAMKYKNYNIDG 382
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 QVIGDLKGLNPLSGSDLIFTSPYMSAVKTVGIIALAEAGVAVGRSFAMFKNYNIDG 360
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 383 NKEMIAIGTMNVVGSFTSCYLLTGPFSRAVNNAGCKTAASNIIMSLAVMLTLLFLTEL 442
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 NKEMIAFGMNI VGSFTCYLLTGPFSRAVNNAGCKTAMSNIVMAIAVMFTLLFLTEL 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 443 FHYTPLVLSAIIIVSAMGLIDYEAAILHLFKVDKDFVVCMSAYIGVFGSVVEIGLVIAI 502
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 421 FHYTPLVLSAIIISAMGLIDYQAAIHLWKVDKDFDLVCMSAYVGVFGSVVEIGLVIAV 480
QY 503 VISLVRLLFIARPRFTVLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYL 562
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 AISIARLLLFVSRPKTAVKGNIPNSMIYRNTQYPSRTPVFGILLILEIDAPIYFANASYL 540
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 563 RERITWIDEEERIKATGETSLQYVIIDMSAVGNIDTSGISMLBEVKKITERRELQVL 622
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 RERIIRWIDEEERVKQSGESSLQYIILDMASVGNIDTSGISMMVEIKKVIDRRALKVL 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 623 VNPVSEVMKLNKSKF-QNHGLGKKWYLTVEEAVGACNFNLASKTNPKKDETEGNNV 680
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 601 SNPKGEVVKLTRSKFTGDHLGKEMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNNV 658

RESULT 3
T51161
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 09-Jul-2004
C:Accession: T51161
R:Comella, P.; Wu, H.J.; Laudie, M.; Berger, C.; Cooke, R.; Delseny, M.; Grellet, P.
Plant Mol. Biol. 41, 687-700, 1999
A:Title: Fine sequence analysis of 60 kb around the Arabidopsis thaliana AtEm1 locus on
A:Reference number: Z24835; MUID:20108326; PMID:10645728
A:Accession: T51161
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-703 <COM>
A:Cross-references: UNIPROT:Q9SV13; EMBL:AF049236; PIDN:AAC14417.1
C:Genetics:
A:Map position: 3
A:Introns: 161/1; 230/2; 287/3; 309/3; 347/3; 386/3; 429/1; 524/3; 588/2; 617/1; 642/3
C:Superfamily: sulfate transport protein

Query Match 75.9%; Score 2638.5; DB 2; Length 703;
Best Local Similarity 75.7%; Pred. No. 4.1e-174;
Matches 499; Conservative 82; Mismatches 75; Indels 3; Gaps 3;

QY 24 MGSVDYEYPLGMNFRVHQ-VEVPPPPFFKSLKYLKETFPPDPLRQFKNKPASKKF 82
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 46 MGTEDTFFQGAELHRRHHTVEAPQPPFLKSLQYSVKETLFPDDPPFQFKQNQASRKF 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 83 MLGLQFFPIFEWAPKTYTQFLKADLIAGITIASLAIPOGISYAKLANLPPILGLYSST 142
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 106 VLGLKYFLPIFEWAPRYNLKFFKSDLIAGITIASLAIPOGISYAKLANLPPILGLYSST 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 143 PPLIYAMGSSRDLAGVTAVGSLMGSLMNAVDNEDPKLYHLAFTATLPAFVFOAA 202
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 166 PPLUYAVLGSSRDLAGVTAVASLLTGAMLSKEVDAEKDPKLYHLAFTATLPAFVLEAS 225
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 203 LGLFRGLIVDFLSHATIIIGFMGGAATVVCLOOLKSLILGLEHFTHGADIISVMRSVFTQT 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 226 LGIFRLGFIIVDFLSHATIVGFMGGAATVVSLOQLKGI FGLKHTDSTDVISMRSVFSQT 285
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 263 HEWRWESAVLGCVFIFLLSTRYFSKRPFRFFWVSAMAPLTSVILGSLLVYFTHAEKHGV 322
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 286 HEWRWESVGLGCCFLFLLSTRYFSIKKPKFFWVAAMAPLTSVILGSLLVYFTHAEKHGV 345
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 323 EVIGELKGLNPPSLNLVPSYMTTAVKTVGVGIISLAEGIAVGRSFAMKYKNYNIDG 382
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 346 QVIGDLKGLNPLSGSDLIFTSPYMSAVKTVGIIALAEAGVAVGRSFAMFKNYNIDG 405
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 383 NKEMIAIGTMNVVGSFTSCYLLTGPFSRAVNNAGCKTAASNIIMSLAVMLTLLFLTEL 442
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 406 NKEMIAFGMNI VGSFTCYLLTGPFSRAVNNAGCKTAMSNIVMAIAVMFTLLFLTEL 465
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 443 FHYTPLVLSAIIIVSAMGLIDYEAAILHLFKVDKDFVVCMSAYIGVFGSVVEIGLVIAI 502
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 466 FHYTPLVLSAIIISAMGLIDYQAAIHLWKVDKDFDLVCMSAYVGVFGSVVEIGLVIAV 525
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 503 VISLVRLLFIARPRFTVLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYL 562
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


Db 526 AISIARLLLFVSRPKTAVKGNIPNSMIYRNTEQYPPSRTVPGILLIILEIDAPIYFANASYL 585
Qy 563 REBITRWIDEEERIIRKATGTSLOYIIDMSAVGNIDTSGISMLBEVKKITERREIQLV 622
Db 586 REIRIIRWIDEEERVKQSGESSLOYIIDMSAVGNIDTSGISMMVEIKKVIDRRAKLVL 645
Qy 623 VNPVSEVMKLNKSKF-QNHGLGKKWIYLTVEEAVGACNFNLASKTNPKKDETEGNNV 680
Db 646 SNPKGEVVKLTRSKFIGDHLGKEMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNV 703

RESULT 4
T48902
sulfate transporter AST12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T48902
R:Takahashi, H.; Sasaki, N.; Kimura, A.; Watanabe, A.; Saito, K.
Plant Physiol. 121, 686, 1999
A:Title: Identification of two leaf-specific sulfate transporters in Arabidopsis thaliana
A:Reference number: Z25001
A:Accession: T48902
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-646 <KAP>
A:Cross-references: UNIPROT:Q9SV13; EMBL:AB012048; PIDN:BA025175.1
A:Experimental source: ecotype Columbia
C:Genetics:
A:Map position: 3
A:Introns: 116/1; 185/2; 242/3; 264/3; 302/3; 329/3; 372/1; 467/3; 531/2; 560/1; 585/3
C:Superfamily: sulfate transport protein

Query Match 74.1%; Score 2575.5; DB 2; Length 646;
Best Local Similarity 74.4%; Pred. No. 8.2e-170;
Matches 490; Conservative 80; Mismatches 74; Indels 15; Gaps 4;

Qy 24 MGSVDYEPYGLMGNFVRHQ-VEVPPPPQPFKSLKYSKTFPPDDPLRQFKKPKASKKF 82
Db 1 MGTEYTFPQGAELHRRHRTVPAOPQPFKSLQYSKVTLPDPDPFPQFKQNASKRF 60
Qy 83 MLGLOPFPIFEWAPKYTFQFLKADLIAGITIASLAIPQGISYAKLANLPPIGLYSSFI 142
Db 61 VLGLKYLPLFEWAPRYNLKFKSDLIAGITIASLAIPQGISYAKLANLPPIGLYSSFV 120
Qy 143 PPIIYAMGSSRDLAGVTGAVGSLMGSLMSNAVDNEDPKLYLHLAFTATLPAFVQAA 202
Db 121 PPLVAVVLGSSRDLAGVTGAVASLLTGAMLSKEVDAEKDPKLYLHLAFTATPAGVLEAS 180
Qy 203 LGLFRGLIYDFLSHATIIQFMGGAATVCLQOLKSLGLEHETHGADIISVMRSVFTQT 262
Db 181 LGIYFRGLIYDFLSHATIIQFMGGAATVWSLOQLKGIYGLKHTDSDTIVSMRSVFSQT 240
Qy 263 HEWRWESAVLGCVFIFFLLSTRYFSKKRPFFWVSAMAPLTSVILGSLLYVYTHAEKHGV 322
Db 241 HEWRWESVGLCGFLPFLSTRIFS KKPFFWVAMAPLTSVILGSLLYVYTHAEKHGV 300
Qy 323 EVIGELKKGLNPFLSLNLFVSPYMTTAVKTGIVVGIISLAEGIAVGRSFAMYNINIDG 382
Db 301 QV-----GSDLIFTSPVMSTAVKTGLITGIIALAEAGVAVGRSFAMFNINIDG 348
Qy 383 NKEMIAIGTNVVGSTFTSCYLTGPFSSRSVAVNAGCKTAASNIIMSLAWMLTLLFLTPL 442
Db 349 NKEMIAFGMNNIVGSEFTSCYLTGPFSSRSVAVNAGCKTAMSNIVMAIVMFTLLFLTPL 408
Qy 443 FHTPLVLSAITSAMGLIDIEAAIHLFKVDKDFVWCMSAYIGVFGSVIGLVI 502
Db 409 FHTPLVLSAITSAMGLIDIOAAIHLWKVDKDFVLCMSAYGVVFGSVIGLVI 468
Qy 503 VISLVRLLFIARPTFVLGNIPNSVIRVYQNAKHVPGMLIILEIDAPIYFANASYL 562
Db 469 AISIARLLLFVSRPKTAVKGNIPNSMIYRNTEQYPPSRTVPGILLIILEIDAPIYFANASYL 528
Qy 563 REBITRWIDEEERIIRKATGTSLOYIIDMSAVGNIDTSGISMLEEVKKITERREIQLV 622

Db 529 REIRIIRWIDEEERVKQSGESSLOYIIDMSAVGNIDTSGISMMVEIKKVIDRRAKLVL 588
Qy 623 VNPVSEVMKLNKSKF-QNHGLGKKWIYLTVEEAVGACNFNLASKTNPKKDETEGNNV 680
Db 589 SNPKGEVVKLTRSKFIGDHLGKEMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNV 646

RESULT 5
T01079
sulfate transport protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01079; T52295
R:Kaplan, N.; Johnson, D.; Schutz, K.; Gnoj, L.; Hoffman, J.; Till, S.; de la Bastide, M.
hi, M.; Martienssen, R.; Chen, E.Y.; Wilson, R.; McCombie, W.R.
Submitted to the EMBL Data Library, November 1998
A:Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.
A:Reference number: Z14248
A:Accession: T01079
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-646 <KAP>
A:Cross-references: UNIPROT:O04289; EMBL:AC002330; NID:g2262135; PID:g2262137
A:Experimental source: cultivar Columbia
R:Takahashi, H.; Sasaki, N.; Kimura, A.; Watanabe, A.; Saito, K.
Plant Physiol. 121, 686, 1999
A:Title: Identification of two leaf-specific sulfate transporters in Arabidopsis thaliana
A:Reference number: Z25001
A:Accession: T52295
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-646 <TAK>
A:Cross-references: EMBL:AB004060; PIDN:BA020282.1
A:Experimental source: strain Columbia
C:Genetics:
A:Map position: 4
A:Introns: 107/1; 176/2; 233/3; 255/3; 293/3; 332/3; 375/1; 470/3; 534/2; 563/1; 588/3
A:Note: T10P11.3
C:Superfamily: sulfate transport protein
C:Keywords: sulfate transport; transmembrane protein

Query Match 66.6%; Score 2314; DB 2; Length 646;
Best Local Similarity 67.2%; Pred. No. 8.7e-152;
Matches 430; Conservative 104; Mismatches 102; Indels 4; Gaps 2;

Qy 42 HQVEVPPPPQPFKSLKYSKTFPPDDPLRQFKNK-PASKKFMGLQFPFPIFEWAPKYT 100
Db 10 HQVEIPPPQPFKSLKNTLINEILFADDPFRIRNESKTSKKTELGLRHVPPILEWARGYS 69
Qy 101 FQFLKADLIAGITIASLAIPQGISYAKLANLPPIGLYSSFIPLIYAMGSSRDLAGVT 160
Db 70 LEYKSDVISGITIASLAIPQGISYAQLANLPPIGLYSSLVPLVYAIMGSSRDLAGVT 129
Qy 161 VAVGSLMGSLMSNAVDNEDPKLYLHLAFTATLPAFVQAAALGLFRGLIYDFLSHATI 220
Db 130 VAVASLLTAAMLGKEVNAVNVNPKLYLHLAFTATFFAGLMQTCGLGLRLGFWVEILSHAAI 189
Qy 221 IQFMGGAATVCLQOLKSLIGLEHETHGADIISVMRSVFTQTHWRWESAVLGCVIFFL 280
Db 190 VQFMGGAATVCLQOLKSLIGLHLHFTHTSDIVTVLRSIFSQSHWRWESVGLCCFLIFL 249
Qy 281 LSTRYFSKKRPFFWVSAMAPLTSVILGSLLYVYTHAEKHGVEVIGELKKGLNPPLSLTNL 340
Db 250 LTTKYISKRPKPLFWISAMSPVSVIFGTIFLYFLHQFHGIQFIGELKKGINPPSITHL 309
Qy 341 VFVSPYMTTAVKTGIVVGIISLAEGIAVGRSFAMYNINIDGKNEMIAIGTNVVGSTFTS 400
Db 310 VFTPPYVLMALKVGIITGVIALAEGIAVGRSFAMYNINIDGKNEMIAIGTNVVGSTFTS 369
Qy 401 CYLTTGPFSSRSVAVNAGCKTAASNIIMSLAWMLTLLFLTPHYTPLVLSAIIYSAML 460
Db 370 CYLTTGPFSSRSVAVNAGCKTALSNNVMAVAVATLLFLTPFFYTPPLVVLSSIIAAML 429

Qy 461 GLIDYEAAIHLFKVDKFDFFVVCMSAYIGVVGSGVEIGLVIAIVISLVRVLLFIARPRTFV 520
Db 430 GLVDYEAAIHLWKDKDFDFVCLISAYLGVVFGTIEIGLISVGSIVMRLVLFVGRPKIYV 489
Qy 521 LGNI PNSVIVRNVEHYQNAKHVPGLMILEIDAPYIFANASYLRIERTWRIDEEERIKAT 580
Db 490 MGNIQNSEIYRNIDHYPOATITRSLLTLHDGPIYFANSTYLDRIGRWIDDEEDKLRTS 549
Qy 581 GETSLOQVVIIDMSAVGNIDTSGISMLBEVKKITERRRELQVLVNPVSEVMKLNKSKFQN 640
Db 550 GDLSLOQVIVLDSAVGNIDTSGISMLBEELNKLICRRELKLVIANPGAEVWKLSKSTFIE 609
Qy 641 HLGKKWTLYLVEEAVGACNFNLRSKNTPKKDETEGWNV 680
Db 610 SIGKERIYLYTVAEEAAACDFMLHTAKPDSVPVE---PNNV 646

RESULT 6

B86365
probable sulphate transporter protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86365
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chan, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Chung, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B86365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-631 <STO>
A;Cross-references: UNIPROT:Q9SKS2; GB:AE005172; NID:g2829902; PIDN:AAC00610.1; GSPDB:GN
A;Genetics:
A;Map position: 1
C;Superfamily: sulfate transport protein

Query Match 54.6%; Score 1898.5; DB 2; Length 631;
Best Local Similarity 57.1%; Pred. No. 3.7e-123;
Matches 353; Conservative 125; Mismatches 139; Indels 1; Gaps 1;

Qy 41 VHQVEVPPQPFKSKLSYLSKETFFDDPLRQFNKNPASKKFMGLGQFFPPIPEWAPKYT 100
Db 3 VHKVAPPHKSTVAKLTKLKETFFDDPLRQFNKNPASKKFMGLGQFFPPIPEWAPKYT 62
Qy 101 FQELKADLIAGITIASIAIQGISYAKLANLPILGLYSSFIPPLIYAMGSSRDIAVGT 160
Db 63 FSLKSDVSVGLTASIAIAIQGISYAKLANLPILGLYSSFIPPLIYAMGSSRDIAVGT 122
Qy 161 VAVGSLIMGMLSNVDPNPDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLSHATI 220
Db 123 VSTASLILGSMLRQVSPVDDVPLFQLAPSSFFAGLFOASGLIIRLGFIDFLSKATL 182
Qy 221 IGFVGGAATVVCLOQLSKILGLEHFTHGADIISVMRSVFQTQTHWRWESAVLGCVIFFL 280
Db 183 IGFVGGAATVVCLOQLSKILGLEHFTHGADIISVMRSVFQTQTHWRWESAVLGCVIFFL 242
Qy 281 LSTRYFSKKEPRPFWSAMAPLTSVILGSLLYVFTAEKHGVEVIGELKKLNPPSLTNL 340
Db 243 LSTRHLSMKKPKLFWWSAGAPLSVISTVLLVFPVFAERHGISVIGKLEGLNPPSWNML 302
Qy 341 VFVSPYMTTAVKTVGVVGIISLAEGIAVGRSFAMYNKYNIDGNKEMIAICTMNVVGSFTS 400
Db 303 QFHGSHLIALVAKTGLVTGIVSLTEGIAVGRTPAALKNYHVDGNKEMIAIGLMNVVGSATS 362
Qy 401 CYLTGTPFSSRVNYNAGCKTAASNIIMSLAVMLTLFLTPFLPHYTPLVVLSAIIVSAML 460

[illegible]

Db 326 TFFVYITRADKGVQIVKHLKGLNPSLSLIIYFSGDYLLKGFRIQVSGHVALTEAVAI 385
Qy 369 GRSPAMYKNYNDGNKEMIAIGTNVVGSTCYLTGTFPSRSVAVNAGCKTAASNIIM 428
Db 386 GRTFAAMKDYQIDGNKEMALGAMNVIGSMSTCVSTGTSFSGRSVAVNPMAGCQTAVSNIIM 445
Qy 429 SLAVMLTLFLTPLFHYTLPVLVLSAIIIVSAMLGLIDYEAAILHFKVDKDFVVCMSYIG 488
Db 446 SIVVLLTLFLTPLFKTPYTNAILAAIINAVIPVDVNAVILFKIDKDFVACMGAPFG 505
Qy 489 VVFGSVEIGLIIAVISVLRVLLFIARPTFVLGNIPNSVIYRNVEHYQNAKHPVGMIL 548
Db 506 VIFVSVEIGLIIAVISVLRVLLFIARPTFVLGNIPNSVIYRNVEHYQNAKHPVGMIL 565
Qy 549 EIDAPIYFANASYLRERITWIDEEERIKATGETSLQYVIIDMSAVGNIDTSGISMLEE 608
Db 566 RVDSAIYFNSVYRRIQRLWLTDEEMVEAARLPRIQFLIEMSPVTDIDTSGIHALED 625
Qy 609 VKKITERRELQVLVNPVSEVMKLNKSKFQNLKWKIYLTVEEAVGACN 659
Db 626 LYKSLQKRDIOQLVLANPFPVINKLVSHFADLIGHDKIFLTVAEAVDSCS 676
RESULT 8
D96809
Protein F28K19.22 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D96809
R;Theologas, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96809
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-719 <STO>
A;Cross-references: UNIPROT:Q9MAX3; GB:AE005173; NID:96573765; PIDN:AAF17685.1; GSPDB:GN
C;Gene: F28K19.22
A;Map position: 1
C;Superfamily: sulfate transport protein
Query Match 50.6%; Score 1758.5; DB 2; Length 719;
Best Local Similarity 49.9%; Pred. No. 2e-113;
Matches 342; Conservative 116; Mismatches 157; Indels 71; Gaps 5;
Qy 42 HQVEVPPQPPFKSLKYSLETFFDDPLRQFNKPKASKFMLGLQFPFPIFWAPKYTF 101
Db 28 HKVGIPEQKNWDFMYTFETTFDDPLRQFNKPKASKFMLGLQFPFPIFWAPKYTF 87
Qy 102 QELKADLIAGITIASLAIPQ-----GISVAKLANLPPIILGLYSSF 141
Db 88 KFRGDLISGLTITASLCIPQVKNLSNCTSTSYLLVLYVDIGYAKLANDPKYGLYSP 147
Qy 142 IPLLAYMMGSSRLDVGTVAVGSLMGSLMNAVDNEDPKLYLHLAFTATLPAQVFOA 201
Db 148 VPPLVACMGSSRDIAIGPVAVVSLGLTGLRAEIDNTPSPDEYLRLAFTATFAGITEA 207
Qy 202 ALGLFRGLIVDFLSHATIIIGFMGGAATVVCLOOLKSKILGLEHFTHGADIISVWRSVFTQ 261
Db 208 ALGFFRGLGIDFLSHAHVAVGFMGGAATVVCLOOLKSKILGLEHFTHGADIISVWRSVFTQ 267
Qy 262 THE-WRWSAVLGCVFIFELLSTRYF-----SKRPP 291
Db 268 AHGWNWVILIGASFLTFLTSLKIIVRHISINKTSKIFLCLDLFLTSLDMLKQKSK 327

Qy 292 RFPWVGAMAPLTSVILGSLLVYPTHAEKHGVE-----VIGELKKGL 332
Db 328 KLFVWPAIAPLISVIVSTFTFVITRADKQGVQIVRSQPLTSLFRFKQFVVVVKHLDOGI 387
Qy 333 NPESLTNLVFPVSPYMTAVKTGIWVGIIISLAEGIAVGRSFAMKYNKINDGNKEMIAIGTM 392
Db 388 NPFSFHLIYFTGDNLAGIRIGVVGWVALT--VTGRTFAAMKDYQIDGNKEMVAGMM 445
Qy 393 NVVGSTCYLTGTFPSRSVAVNAGCKTAASNIIMSLAVMLTLFLTPLFHYTLPVLVLS 452
Db 446 NVVGSMSSCVATGTSFSGRSVAVNPMAGCQTAVSNIIMSVLLTLLFLTPFKYTPNAILA 505
Qy 453 AIIIVSAMLGLIDYEAAILHFKVDKDFVVCMSYIYGVFGSVEIGLIIAVISVLRVLLF 512
Db 506 AIIINAVIPLIDITQAAAILFKVDKDFIACIAGFFGVIFVSVBEIGLIIAVISFAKILLQ 565
Qy 513 IARPTFVLGNIPNSVIYRNVEHYQNAKHPVGMILLEIDAPIYFANASYLRERITWIDE 572
Db 566 VTRPRTAVLGNIPRTSVYRNIOQYPEATVPGVLTIRVDSAIYFNSVYRRIQRLWHE 625
Qy 573 EERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKK 632
Db 626 EEEKVKAASLPRIQFLIEMSPVTDIDTSGIHALEDLYKSLQKRDIOQLVLANPGLVIGK 685
Qy 633 LNKSKFQNLKWKIYLTVEEAVGAC 658
Db 686 LHLSHFADMLGDQNIYLTVADEAVAC 711
RESULT 9
T00946
Probable sulfate transport protein T3F12.7 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00946
R;Gnoj, L.; Huang, B.N.; Habermann, K.; Hameed, A.; Hasegawa, A.; Jensen, K.; Schutz, K.;
R.; McCombie, W.R.
submitted to the EMBL Data Library, October 1997
A;Description: Arabidopsis thaliana BAC T3F12 from chromosome IV.
A;Reference number: Z14210
A;Accession: T00946
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-649 <GNO>
A;Cross-references: UNIPROT:Q9SAY1; EMBL:AC002983; NID:G2443899; PID:G2565006
C;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Intons: 102/3; 119/1; 246/3; 306/3; 345/3; 388/1; 483/3; 547/2; 576/1; 601/3
A;Note: T3F12.7
C;Superfamily: sulfate transport protein
C;Keywords: sulfate transport; transmembrane protein
Query Match 50.4%; Score 1752; DB 2; Length 649;
Best Local Similarity 52.4%; Pred. No. 4.9e-113;
Matches 333; Conservative 124; Mismatches 177; Indels 2; Gaps 2;
Qy 34 GWNFVHVQ-VEVPPQPPFKSLKYSLETFFDDPLRQFNKPKASKFMLGLQFPFPI 92
Db 14 GARNPPVQRVLAPPKAGLLDKSVBEETFFHDAPLDFKQGTAKKALLGQIAVFPI 73
Qy 93 FEWAPKYTFQFLKADLIAGITIASLAIPQGISVAKLANLPPIILGLYSSFPPIIYAMGMS 152
Db 74 IGWAREYTLRKFRGDLIAGITIASLCIPQDIGYAKLANDPKYGLYSSFPPIIYAMGMS 133
Qy 153 SRDLAVGTAVGSLMGSLMNAVDNEDPKLYLHLAFTATLPAQVFOAALGLFRGLIV 212
Db 134 SRDIAIGPVAVVSLVGLTLCQAVIDPKNPKEDYLRVFTATFPAQIFAGLGLFRGLI 193
Qy 213 DFLSHATIIIGFMGGAATVVCLOOLKSKILGLEHFTHGADIISVWRSVFTQT-HEMWRESAV 271
Db 194 DFLSHAHVAVGFMGGAATVVCLOOLKSKILGLEHFTHGADIISVWRSVFTQT-HEMWRESAV 253

QY	379	NIDCNKEMIAIGTMNVVGSFTSCYLTTPGFSRANVYAGCKTAASNIIMSLAVMLTLF	438
Db	381	ALDGNKEMWAMGTWNI VGSLSVCVTTGSPFSRANVYAGCKTAVSNVMSIVLLTLV	440
QY	439	LTLFPHYTPVLVLSAIVSAMGLIDYEAIAHLFKVDKDFVVCMSAYIGVFGSVEIGL	498
Db	441	ITPLEKFTPNANVLASIIIAAVNLVNIEMVLLWKIDKDFVACMGAFGVIKFSVEIGL	500
QY	499	VIAIVSLVRLVLFARPRFVLGNTPNSVIYRNEHYQNAKHVPQMLLILEIDAPYFAN	558
Db	501	LIAVAISFAKILLQVTRPTAVLGKIPGTSVYRNIQQYPKAAQIPGMLIIRVDSAIYFSN	560
QY	559	ASYLREIRTWIDEEERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERREL	618
Db	561	SNVIKERILRWLIDEGAQRTESELPEIQHLITEMSVPVDIDTSGIHAFEELYKTLQREV	620
QY	619	QLVLVNPVSVEMKLNKSKFQNHGKKWYILVVEEAV	655
Db	621	QLILANPGVPVIEKLHASKLTTELIGEDIPLTVADAV	657
RESULT 14			
T50022			
sulfate transporter - Arabidopsis thaliana			
N;Alternate names: protein T31P16.170			
C;Species: Arabidopsis thaliana (mouse-ear cress)			
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004			
C;Accession: T50022			
R;Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Kalicki, J.; Wohldmann, P.; Smith, P.M.; Ealing, P.M.; Hawkesford, M.J.; Clarkson, D.T.			
submitted to the Protein Sequence Database, May 2000			
A;Reference number: Z25027			
A;Accession: T50022			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-677 <BEV>			
A;Cross-references: UNIPROT:O04722; EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.170			
A;Experimental source: cultivar Columbia; BAC clone T31P16			
C;Genetics:			
A;Gene: ATSP:T31P16.170			
A;Map position: 5			
A;Introns: 132/3; 149/1; 298/3; 375/3; 418/1; 513/3; 578/2; 609/1; 634/3			
C;Superfamily: sulfate transport protein			
Query Match 43.4%; Score 1507; DB 2; Length 677;			
Best Local Similarity 48.2%; Pred. No. 4e-96;			
Matches 300; Conservative 128; Mismatches 183; Indels 12; Gaps 8;			
QY	48	PPQPFKSLKYSUKETFFPDPLRQFN---KPASKKFMGLQFPFPPIFEWAPKYYTFQL	104
Db	59	PPSPWHE-LKRQVKGSL--TKAKFKSLQKQFPKQILSVLQAIPIFGWCRNYKLTWF	115
QY	105	KADLIAGITIASIAIQGISYAKLANLPILGLYSSFIPLIYAMGSSRDLAGVTVAG	164
Db	116	KNDLMAGLTLASLCIPQSIGYATLAKLDPOQGLYTSWPPPLIYALMGTSRETAIGPVAV	175
QY	165	SLIIMGSLNVAVDPNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLSHATIIGFM	224
Db	176	SLIISMLQKLIIDPETHDPLGYKKLVITTTFFAGIFQASGLFRLGLVDFLSHAALVGM	235
QY	225	GGAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQ--THEWRWESAVLGCVFIFFLST	283
Db	236	GGAAIVIGLQQLKGLGIGNFTNTNTDIVSVLVRWSCQQQNSPHTFILGCSFLSPILIT	295
QY	284	RYFSKKRPRFFWVSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGILNPDLNLVVF	343
Db	296	RFTGKKYKGLFMDLPAFLVAVVSTLMVFLTKADEHGKVTVRHIIKGLNPMISIQOLDEN	355
QY	344	SPYMTTAVKTGIVVGIISLAEGIAGVRSFAMYNKYNIDGNKEMIAIGTMNVVGSFTSCYL	403
Db	356	TPHLGQIAKIGLIIAIVLAEAVGRSPAGIKGYRLDGNKEMVAGFNNVLGFSFTSCYA	415
QY	404	TTGPFSSRANVYAGCKTAASNIIMSLAVMLTLF	463

Db	416	ATGSPSRFANFAAGCETAMSNVMAVTVFALECLTRLTYTPIAILASIIISALPGLI	475
QY	464	DYEAIAHLFKVDKDFVVCMSAYIGVFGSVEIGLVAIVSLVRLVFIAPRPTFVLGN	523
Db	476	NINEAIIHWKVDKDFLADIGAPFGLFASVEIGLVAIVSFAKIIISIRPIETLGR	535
QY	524	IPNSVIYRNEHYQNAKHVPQMLLILEI-DAPYFANASYLRERITRWID--BEEERIKAT	580
Db	536	MPGTDFTDTNQYPMWTKTPGVLIIRVKSALCFANASSIEERIMGWDEBEEENTKSN	595
QY	581	GTSIQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKLNKSKPON	640
Db	596	AKRKILFVVLDMSSLINVDTSITALLELHNKLIKRTGVELVIVNPKQVIHKLNOAKPVD	655
QY	641	HLGKKWYILVVEEAVCAACNFNL	663
Db	656	RIGGK-VYLTIGEALDAC-FGLK	676
RESULT 15			
S51765			
sulfate transport protein 3, low affinity - Stylosanthes hamata			
C;Species: Stylosanthes hamata			
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004			
C;Accession: S51765			
R;Smith, P.W.; Ealing, P.M.; Hawkesford, M.J.; Clarkson, D.T.			
submitted to the EMBL Data Library, November 1994			
A;Description: Cloning and expression of cDNAs that encode sulphate transporters from pl			
A;Reference number: S51765			
A;Accession: S51765			
A;Molecule type: mRNA			
A;Residues: 1-644 <SMI>			
A;Cross-references: UNIPROT:P53393; EMBL:X82454; NID:g607187; PIDN:CAA57831.1; PID:g60718			
C;Genetics:			
A;Gene: SHST3			
C;Superfamily: sulfate transport protein			
C;Keywords: sulfate transport			
Query Match 43.3%; Score 1503; DB 2; Length 644;			
Best Local Similarity 47.0%; Pred. No. 7.1e-96;			
Matches 301; Conservative 124; Mismatches 193; Indels 22; Gaps 6;			
QY	33	LGWNNFVRVHQ--VEVPPPPQPFKSLKYSUKETFFPDPLRQFNKPASKKFMGLQF--	88
Db	4	LGTEQFSRSQWVLSNPPLTKKFLGLKD-----NKFFTSSSSKKETRAVSFLA	55
QY	89	-PPIFEWAPKYYTFQLKADLIAGITIASIAIQGISYAKLANLPILGLYSSFIPLIY	147
Db	56	SUFPILSMIRTVSATKFKDDLLSGULTASLSIPQSIGYANLAKLDPOQGLYTSVIPPVY	115
QY	148	ANMGSSRDLAGVTVAGSILMGSLNVAVDPNEDPKLYLHLAFTATLFAGVFOAALGLPR	207
Db	116	ALMGSSRETAIGPVAVVSMILSSLPVKVIDPDAHENDYRNLVFTVTLFAGIFQAFGLUR	175
QY	208	LGILVDFLSHATIIIGFMGGAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQTHE---	264
Db	176	LGFLVDFLSHAALVGMAGAAIVIGLQQLKGLGLTHFTTKTDAVAVLSVYTSLHQQT	235
QY	265	-----NRWESAVLGCVFIFPLLSLTRVFSKKRPRFFWVSAMAPLTSVILGSLVYFTHAEKH	320
Db	236	SSENWSPLNFVIGCSFLIFLAAARPIGRNKKFFLPAIPALPLSVILSTLIVFLSKGDKH	295
QY	321	GVEVIGELKKGILNPDLNLVVFVPMYTTAVKTGIVVGIISLAEGIAGVRSFAMYNKYN	380
Db	296	GVNIITKHVGGGLNPSVSHKQLNGHPVGOAKIGIHSAIIALTEAIAVGRSPANIKGHL	355
QY	381	DGNKEMIAIGTMNVVGSFTSCYLTTPGFSRANVYAGCKTAASNIIMSLAVMLTLF	440
Db	356	DGNKEMLAWGCNIIAGSLTSCVSTGSPSRFANVYAGCKTAVSNVMAVTVLLCLELFT	415
QY	441	PLFHYTPLVLSAIVSAMGLIDYEAIAHLFKVDKDFVVCMSAYIGVFGSVEIGLVI	500

Db 416 RLLYTPMAILASIIISALPGLIDIGEAYHIWKVDKFDLACLGAFFGVLFVSTIEIGLLI 475
Qy 501 AIVISVLRLVLLFIARPTTFVLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDA-PIYPANA 559
Db 476 ALSISFAKILLQAIRPGVEVLGRIPPTTEAYCDVAQYPMVTTTGGILVIRISSGSLCFANA 535
Qy 560 SYLREIRITRWI-DEEERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERREL 618
Db 536 GFVREIRILKWVEDEEQDNIEEAAKGRVQAIIDMTDLTNVDTSGILALEELHKKLLSRGV 595
Qy 619 QLVLVNVPSEVMKGLNKSKFQNLGKKWYILTVEEAVGAC 658
Db 596 ELAMVNPWEVIHKLKVANFVDKIGKERVFLTVAEAVDAC 635

Search completed: August 31, 2005, 11:51:58
Job time : 37 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2005, 06:33:57 ; Search time 141 Seconds
(without alignments)
2469.602 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLRNTI.....NLRASKTNPKKDETEGNNV 680

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2641.5	76.0	658	1	ST31_ARATH	Q98v13 arabidopsis
2	2596.5	74.7	658	2	Q70GJ1	Q70GJ1 brassica na
3	2326.5	66.9	637	2	Q8H7X7	Q8H7X7 oryza sativ
4	2314	66.6	646	1	ST32_ARATH	Q04289 arabidopsis
5	2274.5	65.5	650	2	Q70BL1	Q70BL1 brassica ol
6	2057	59.2	659	2	Q7XE18	Q7XE18 oryza sativ
7	1898.5	54.6	631	1	ST33_ARATH	Q98x82 arabidopsis
8	1880.5	54.1	656	2	Q9AT47	Q9AT47 lycopersico
9	1875	54.0	653	1	ST34_ARATH	Q91w86 arabidopsis
10	1858	53.5	656	2	Q6Z9V1	Q6Z9V1 oryza sativ
11	1850.5	53.3	658	2	Q9AT12	Q9AT12 zea mays (m
12	1841.5	53.0	653	1	ST12_ARATH	Q9max3 arabidopsis
13	1836.5	52.8	629	2	Q7X6T0	Q7X6T0 oryza sativ
14	1830.5	52.7	656	2	Q6ZXB7	Q6ZXB7 brassica ol
15	1811.5	52.1	666	2	Q942E2	Q942E2 oryza sativ
16	1810.5	52.1	655	2	Q93X17	Q93X17 brassica na
17	1809.5	52.1	635	2	Q6ZZ94	Q6ZZ94 brassica ol
18	1807.5	52.0	656	1	ST13_ARATH	Q9fep7 arabidopsis
19	1796.5	51.7	651	2	Q94FJ2	Q94FJ2 lycopersico
20	1791.5	51.6	657	2	Q9AT48	Q9AT48 lycopersico
21	1787	51.4	634	1	ST35_ARATH	Q94lw6 arabidopsis
22	1777.5	51.2	662	2	Q8H7X3	Q8H7X3 oryza sativ
23	1773.5	51.0	662	2	Q8RUC3	Q8RUC3 oryza sativ
24	1770.5	50.9	648	2	Q93YK6	Q93YK6 brassica na
25	1767.5	50.9	657	2	Q9FPQ4	Q9FPQ4 solanum tub
26	1760.5	50.7	660	2	Q43482	Q43482 hordeum vul
27	1760.5	50.7	660	2	Q84XE9	Q84XE9 hordeum vul
28	1752	50.4	649	1	ST11_ARATH	Q9say1 arabidopsis
29	1749.5	50.3	660	2	Q40008	Q40008 hordeum vul
30	1745.5	50.2	662	2	Q8H0K3	Q8H0K3 triticum ae
31	1738.5	50.0	662	2	Q8H2D8	Q8H2D8 aegilops ep

32	1738.5	50.0	662	2	Q9XGB5	Q9XGB5 aegilops ta
33	1735.5	49.9	662	1	SUT2_STVHA	P53392 stylosanthe
34	1725.5	49.7	655	2	Q8H0J8	Q8H0J8 triticum ur
35	1723.5	49.6	660	2	O04001	O04001 sporobolus
36	1722.5	49.6	667	1	SUT1_STVHA	P53391 stylosanthe
37	1720.5	49.5	662	2	Q8H0K2	Q8H0K2 triticum ae
38	1714.5	49.3	655	2	Q8H0K6	Q8H0K6 triticum ae
39	1710	49.2	655	2	Q8H0K4	Q8H0K4 triticum ae
40	1697	48.8	655	2	Q8H0D7	Q8H0D7 aegilops sp
41	1690.5	48.6	655	2	Q8H0K5	Q8H0K5 triticum ae
42	1680	48.3	652	2	Q8H7X4	Q8H7X4 oryza sativ
43	1666.5	48.0	666	2	Q8H0J9	Q8H0J9 triticum ur
44	1653	47.6	649	2	Q9XGB6	Q9XGB6 aegilops ta
45	1563.5	45.0	659	2	Q8LR58	Q8LR58 oryza sativ

ALIGNMENTS

RESULT 1
ST31_ARATH STANDARD; PRT; 658 AA.
ID ST31_ARATH STANDARD; PRT; 658 AA.
AC Q9SVI3; Q23250; O64435; O65025;
DT 10-OCT-2003 (Rel. 42; Created)
DT 10-OCT-2003 (Rel. 42; Last sequence update)
DT 25-OCT-2004 (Rel. 45; Last annotation update)
DE Sulfate transporter 3.1 (AST12) (AtST1).
GN Name=SULTR3;1; Synonyms=ST1; OrderedLocusNames=At3g51895;
ORFNames=F4F15.10, ATBM1.15;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta; TISSUE=Aerial parts;
RA Yamaguchi Y., Nakamura T., Harada E., Koizumi N., Sano H.;
RT "Isolation and characterization of a cDNA encoding a sulfate transporter from Arabidopsis thaliana.";
RL (er) Plant Gene Register PGR97-051.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Takahashi H., Sasakura N., Kimura A., Watanabe A., Saito K.;
RT "Identification of two leaf-specific sulfate transporters in Arabidopsis thaliana.";
RL (er) Plant Gene Register PGR99-154.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20108326; PubMed=10645728; DOI=10.1023/A:1006395324818;
RA Comella P., Wu H.-J., Laudie M., Berger C., Cooke R., Delseeny M., Grellet P.;
RT "Fine sequence analysis of 60 kb around the Arabidopsis thaliana AtEm1 locus on chromosome III.";
RL Plant Mol. Biol. 41:687-700 (1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713; DOI=10.1038/35048706;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unsel M., Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B., Delseeny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P., De Simone V., Choiane N., Artiguenave F., Robert C., Brottier P., Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V., Wurmbach E., Drzonek H., Erflie H., Jordan N., Bangert S., Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G., Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simonati B., Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G., Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J., Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,

RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA de Haan M., Maarsee A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argirio A., Flores M., Liguori R., Vitale D.,
RA Mannhaug G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walte A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldbiyuni T.V.,
RA Preus D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idegawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.P.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koeseena E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome.";
RL Science 302:842-846(2003).
CC -!- FUNCTION: H(+)/sulfate cotransporter that may play a role in the
CC regulation of sulfate assimilation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Expressed only in leaves.
CC -!- INDUCTION: By nitrogen starvation, but not by sulfate starvation.
CC -!- SIMILARITY: Belongs to the SLC26A/sulf transporter (TC 2.A.53)
CC family.
CC -!- SIMILARITY: Contains 1 STAS domain.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D89631; BAA21657.1; -;
DR EMBL; AB012048; BAA25175.1; ALT SEQ.
DR EMBL; AF049236; AAC14417.1; ALT_INT.
DR EMBL; AL049711; CAB43310.1; -;
DR EMBL; AY099873; AAM20724.1; -;
DR PIR; T48901; T48901.
DR PIR; T48902; T48902.
DR PIR; T49069; T49069.
DR PIR; T51161; T51161.
DR GeneFarm; 2359; 189.
DR InterPro; IPR02645; STAS.
DR InterPro; IPR001902; Sulph_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR ProDom; PD005210; Antisig_antgnt; 1.
DR TIGRFAMs; TIGR00815; sulf; 1.
DR PROSITE; PS01130; SLC26A; FALSE_NEG.
DR PROSITE; P550801; STAS; 1.

Multigene family; Sulfate transport; Symport; Transmembrane;
KW DOMAIN 1 85 Cytoplasmic (Potential).
KW TRANSPORT. 86 Potential.
FT TRANSMEM 106 Extracellular (Potential).
FT DOMAIN 107 Potential.
FT TRANSMEM 109 Potential.
FT DOMAIN 129 Potential.
FT TRANSMEM 130 Potential.
FT DOMAIN 133 Potential.
FT TRANSMEM 134 Potential.
FT DOMAIN 154 Potential.
FT TRANSMEM 155 Potential.
FT DOMAIN 163 Potential.
FT TRANSMEM 164 Potential.
FT DOMAIN 184 Potential.
FT TRANSMEM 185 Potential.
FT DOMAIN 186 Potential.
FT TRANSMEM 207 Potential.
FT DOMAIN 207 Potential.
FT TRANSMEM 245 Potential.
FT TRANSMEM 246 Potential.
FT DOMAIN 267 Potential.
FT TRANSMEM 271 Potential.
FT DOMAIN 272 Potential.
FT TRANSMEM 292 Potential.
FT DOMAIN 293 Potential.
FT TRANSMEM 332 Potential.
FT DOMAIN 333 Potential.
FT TRANSMEM 354 Potential.
FT DOMAIN 363 Potential.
FT TRANSMEM 364 Potential.
FT DOMAIN 384 Potential.
FT TRANSMEM 395 Potential.
FT TRANSMEM 398 Potential.
FT DOMAIN 425 Potential.
FT TRANSMEM 426 Potential.
FT DOMAIN 447 Potential.
FT TRANSMEM 464 Potential.
FT DOMAIN 465 Potential.
FT TRANSMEM 485 Potential.
FT DOMAIN 486 Potential.
FT TRANSMEM 533 Potential.
FT DOMAIN 533 STAS.
FT TRANSMEM 533 STAS.
FT CONFLICT 208 V -> L (in Ref. 3).
FT CONFLICT 344 V -> I (in Ref. 1).
SQ SEQUENCE 658 AA; 72748 MW; ED1860A3467FB577 CRC64;
Query Match 76.0%; Score 2641.5; DB 1; Length 658;
Best Local Similarity 75.9%; Pred. No. 2.2e-165; Indels 3; Gaps 3;
Matches 500; Conservative 81; Mismatches 75;
QY 24 MGSVDYEPGLMNNFVRHQ-VEVPPQPPFKSLKYLKSTETFPDPLQFQKPKASKKF 82
DB 1 MGTEDYTFQGAELHRRHTVEAQPQPFKSLQSVKTELFPDPPFPQFKNQASKKF 60
QY 83 MGLGLOFPPIFWAPKYTFQFLKADLIAGITITASIAIPOGISYAKLANLPPILGYSSFI 142
DB 61 VLGLKYFLPIFWAPRYNLKFKSGLIAGITITASIAIPOGISYAKLANLPPILGYSSSV 120
QY 143 PPLIYAMGSSRDLAGVTGAVGSLMGSLNAPDNPDKLYLHLAFTATLPAFVQAA 202
DB 121 PPLVAVLGSSRDLAGVTGAVASLITGMLSEVDAEKDPLYLHLAFTATFAGVLEAS 180
QY 203 LGLFRLGLIVDFLSHATIGFMGGAATVVCLOOLKSLILGLEHFTGADIISVMRSVFTQT 262
DB 181 LGIFRLGLIVDFLSHATIGFMGGAATVVCLOOLKSLILGLEHFTGADIISVMRSVFTQT 240
QY 263 HEMRNESAVLGCVFIFELLSTRTYFSKPRPFVWSAMAPLTSVILGSLIVYTHAEKHGV 322
DB 241 HEWRWESVGLGCGFLFLLSTRTYFSKPRPFVWSAMAPLTSVILGSLIVYTHAEKHGV 300
QY 323 EVIGELKKGLNPPSLTNLNVFVSPYMTAVKTIIVGVIITSLAEGIAVGRSFAMKYNIDG 382
DB 301 QVIGLKKGLNPLSGDLIFTSPYMTAVKTIIVGVIITSLAEGIAVGRSFAMKYNIDG 360
QY 383 NKEMIAIGTMNVVGSFTCYLTGTPFSRNVNAGCKTAAASNIIIMSLAVMLTLFLTPL 442
DB 361 NKEMIAFGMMNIVGSGFTCYLTGTPFSRNVNAGCKTAAASNIIIMSLAVMLTLFLTPL 420
QY 443 FHYTFLVLSAIIIVSAMGLIDYEAHILFKVDKDFVVCMSAYIGVYVGSVEIGIIVAI 502
DB 421 FHYTFLVLSAIIIVSAMGLIDYEAHILFKVDKDFVVCMSAYIGVYVGSVEIGIIVAI 480
QY 503 VTSVLRLVLLFIARPTFLVGNIPNSVYVRNVNHYQNAKHVPGMLILEIDAPIYFANASYL 562
DB 481 AISIARLLLFVSRPKTAVKGNIPNSVYVRNVNHYQNAKHVPGMLILEIDAPIYFANASYL 540
QY 563 RERITRWIDEERIKATGETSLQVVIDMSAVGNIDTSGISMLBEVKKITERRELQVLV 622

```
Db 541 REIIIRWIDEERVRQSGESSLOYIILDSAVGNIDTSGISMMVEIKKYIDRRALKVL 600
QY 623 VNPVSEVMKLNKSKF-QNHGLKKWLYLTVEEAVGACNFNLRASTNPKKDETEGNVN 680
Db 601 SNPKGEVVKLTTSKFGIDHGLKEMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNVN 658

RESULT 2
Q70GL1
ID Q70GL1 PRELIMINARY; PRT; 658 AA.
AC Q70GL1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Sulphate transporter.
GN Name=bst3.1;
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RA Buchner P., Stuijver E.E., Hawkesford M.J., de Kok L.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ581745; CAB46442.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transpt.
DR Pfam; PF01740; STAS; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS50801; STAS; 1.
SQ SEQUENCE 658 AA; 72734 MW; 2EF39BD46340E17E CRC64;

Query Match 74.4%; Score 2596.5; DB 2; Length 658;
Best Local Similarity 75.0%; Pred. No. 2e-162;
Matches 494; Conservative 80; Mismatches 82; Indels 3; Gaps 3;

QY 24 MGSVDVEYPLGMN-FERHQVVEVPPQPPFKSLKSLKXKTEFPDDPLRQPKNKPASKFP 82
Db 1 MGTEDTTFQGAEEPHRRHHAEPDPQPKSLQSVKTELPDDPFQFKNQTTSRQV 60
QY 83 MLGLQFPFFIFEWAPKTYFOFLKADLIAGTIIASLAIPOGISYAKLANLPPIGLYSSFI 142
Db 61 VLGLKFLPLEWAPLYNFKLKSDDLIIAGTIIASLAIPOGISYAKLANLPPIGLYSSFV 120
QY 143 PPIYAMGSSRDLAGVTGAVGSLMGSLMSNAVDNEDPKLYLHLAFTATLPAFVFOAA 202
Db 121 PPLVFAVLGSSKDLAVGTAVGSLLTGAMLSKEVDAEKDPKLYLHLAFTATFPAGVLEAS 180
QY 203 LGLFRGLIVDFLSHATIGFMGGAATVCLQKLSILGLEHHTHADIIISVNRVFTQT 262
Db 181 LGLFRGLIVDFLSHATIGFMGGAATVSLQKLGIFGLKHFTEATDVISNRSVSQT 240
QY 263 HEWRWESAVLGCYFIFPFLSLSTRYFKRPRFFWVSAMAPLTSVILGSLLVYFTHSKGV 322
Db 241 HQWRWESVGLGCCFLPFLSLSTRYFKRPPKPFVWAAWAPLTSVILGSLLVYFTHSRHG 300
QY 323 EVIGELKKGILNPPSLTNLVSPYMTTAVKTVGIWVGIIISLAEGIAVGRSFAMKYKNYNDG 382
Db 301 QVIGDLKKGILNPLSVSDLVFTSPYMTALKTGLITGLTAEGLAVGRSFAMKYKNYNDG 360
QY 383 NKEMIAIGTNVVGSTFSTCYLTGTPSRSAVNTNAGCKTAASNIIMSLAVMLTLLFLTPL 442
Db 361 NKEMIAFGMNIIVGSTFSTCYLTGTPSRSAVNTNAGCKTAVSINVMAIAMVMTLLFHTPF 420
QY 443 FHYTPLVLSAIVSAMGLIIDEAAHLFKVDKDFVVCMSAYIGVFGSVEIGLVAVI 502
```

```
Db 421 FHYTPLVLSIIIMVAMLGLIDYQAAHLWKVDKDFPVCMSAYFGVFGSVEIGLVAV 480
QY 503 VTSVLRVLFIARPTFVLGNIPNSVIYRNVEHYQNAKHVPGLMILILEIDAPIYFANASYL 562
Db 481 VISIARLLLFVSRPRTAVKGNIPNTMIYRNTDQYPYSRIVPGLLILEIDAPIYFANAGYL 540
QY 563 RERITRWIDEERIKATGETSIQYVIDMSAVGNIDTSGISMLREVKKITRERELQVL 622
Db 541 RERITRWIDEEDRIKASGNSLQYVIDMSAVGNIDTSGISMMEEIKKIMDRRELKVL 600
QY 623 VNPVSEVMKLNKSKF-QNHGLKKWLYLTVEEAVGACNFNLRASTNPKKDETEGNVN 680
Db 601 ANPKGEVVKLTTSKFGIDHGLKEMFLTVGEAVEACSFMLHTSKTEPASKE-EPNVN 658

RESULT 3
Q8H7X7
ID Q8H7X7 PRELIMINARY; PRT; 637 AA.
AC Q8H7X7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative sulfate transporter AtST1.
GN Name=OJ1607A12.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Sasaki C.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC105729; AAN06871.1; -.
DR Gramene; Q8H7X7; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR003658; Antisig_antgnt.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR ProDom; PD005210; Antisig_antgnt; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS50801; STAS; 1.
SQ SEQUENCE 637 AA; 69159 MW; DD09D98D86F35059 CRC64;
```

Query Match 66.9%; Score 2326.5; DB 2; Length 637;
Best Local Similarity 70.7%; Pred. No. 1.1e-144;
Matches 440; Conservative 82; Mismatches 99; Indels 1; Gaps 1;

```
QY 46 VPPQPPFKSLKSLKXKTEFPDDPLRQF-KNKPASKKFMGLQFFPPIFEWAPKYTFQEL 104
Db 1 MPAAPKPLETGLGNMKETFLPDDPFVRVRRCGCGRAAALRYVFPFWEWAPSYLTGL 60
QY 105 KADLTAGITIASLAIPOGISYAKLANLPPIGLYSSFIPPLIYAMGSSRDLAGVTGAVG 164
Db 61 KSDLTAGITIASLAIPOGISYAKLANLPVILGLYSSFVPLIYAMGSSRDLAGVTGAVA 120
QY 165 SLLMGSLMSNAVDNEDPKLYLHLAFTATLPAFVFOAALGLFRGLIIVDFLSHATIGRM 224
Db 121 SLLIGSLMSEESAEDPALYLHVALTATFFAGVFOALLGLVLRGLFIVDFLSHATIGRM 180
QY 225 GGAATVVCLOQLKSLGLEHHTHADIIISVMSVFTQTHWESAVLGCVFIFFLLSSTR 284
Db 181 GGAATVVCLOQLKSGFGLDHTTATDLVSVMSVFSQTHLWRWESVVMCGFLFLITR 240
QY 285 YFSKRRPRFFWVSAMAPLTSVILGSLLVYFTHAEKHGVEVIGELKKNPPSLTNLVFVS 344
Db 241 FFSKRRPRFFWVSAMAPLTSVILGSLLVYFTHAENHGIQVIGYKGLNPPSPSTSLNFS 300
```

QY	345	PYMTAVKTIQIVVGIISLAEGIAVGRSFAMFYKNYNIDGNKEMIAIGTMNVVGSFTSCYLT	404	RA	Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,	
Db	301	PYMLALKGTIIGVIALAEGIAVGRSFAMFYKNYHIDGNKEMIAFGTMIVGSLTSCYLT	360	RA	Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,	
QY	405	TGPFSSAVNNAGCKTAASNIIMSIAVMTLLFLPFLPHYPLVVLVSAAIIVSAMGLID	464	RA	Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,	
Db	361	TGPFSSAVNNAGCKTAMSNVIMSVAVMTLLFLPFLPHYPLVVLVSAAIIVSAMGLID	420	RA	Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,	
QY	465	YEAATHLKVDFVVCVSAYIGVVGVSVEIGLVAIVLSVLVLLFTARPTTVLGNL	524	RA	Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,	
Db	421	YPAAVHLWQVDFVVCVGLGVVGSVEIGLVAVGSLVRLVLFVARPTTVLGNL	480	RA	Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,	
QY	525	PNSVIYRNVHYQNAKHPQMLLEIDAPIYFANASYLRITRWDDEEERIKATGETS	584	RA	Fishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,	
Db	481	PNTMIYRMDQYTAQRFVGLVLRVDSPIYFNASYLRERTARWIDDEECKEKGEMG	540	RA	Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,	
QY	585	LQVIIDMSAVGNIDTSGISMLSEVKKITERRELQVLVNPVSEVMKLNKSKFQNLGK	644	RA	Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,	
Db	541	IQVVLDMGAVGSDTSGTSMLEKTLDRGLQVLANPGSEIMKKLDSKVLKLEIGH	600	RA	Sekhon L., Murray J., Sheet P., Cordes M., Abu-Threideh J.,	
QY	645	KWYLTVEEAVGACNENLRASK	666	RA	Stonking T., Kalicki J., Graves T., Harmon G., Edwards J.,	
Db	601	EWIPTVGEAECDFVMSHQ	622	RA	Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,	
RESULT 4					RA	Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
ID	ST32	ARATH	STANDARD;	PRT;	646	AA.
AC	004289;					
DT	10-OCT-2003	(Rel. 42, Created)				
DT	10-OCT-2003	(Rel. 42, Last sequence update)				
DT	25-OCT-2004	(Rel. 45, Last annotation update)				
DE	Sulfate transporter 3.2 (AS777).					
GN	Names-SUTR3.2; OrderedLocusNames=At4g02700; ORFNames=TI0P11.3;					
OS	Arabidopsis thaliana (Mouse-ear cress).					
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;					
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.					
OX	NCBI_TaxID=3702;					
RL	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=cev. Columbia;					
RA	Takahashi H., Sasakura N., Kimura A., Watanabe A., Saito K.;					
RT	"Identification of two leaf-specific sulfate transporters in					
RL	Arabidopsis thaliana."					
RL	{er} Plant Gene Register	PGR99-154.				
RP	SEQUENCE FROM N.A.					
RC	STRAIN=cev. Columbia;					
RA	Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,					
RA	Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,					
RA	Harris B., Anson W., Brandt P., Griwell L.A., Rieger M.,					
RA	Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,					
RA	Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,					
RA	Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,					
RA	Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,					
RA	Langham S.-A., McCullagh B., Bilham L., Robben J.,					
RA	Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,					
RA	Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,					
RA	Weitzenecker T., Bothe G., Rampsperger U., Hilbert H., Braun M.,					
RA	Holzer E., Brandt A., Peters S., van Staveren M., Dirks W.,					
RA	Mooljhan P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,					
RA	Berkeiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,					
RA	De Keyser A., Buyschaert C., Glens J., Villarroel R., De Clercq R.,					
RA	Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,					
RA	Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,					
RA	Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,					
RA	Borkova D., Bloecker H., Schaefer M., Grimm M., Loehner T.-H.,					
RA	Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,					
RA	Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,					
RA	Neumann S., Argiriou A., Vitale D., Liguori R., Piravanti E.,					

RA	Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,	1	76	Cytoplasmic (Potential).
RA	Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,	77	97	Potential.
RA	Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,	98	99	Extracellular (Potential).
RA	Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,	100	120	Potential.
RA	Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,	121	124	Cytoplasmic (Potential).
RA	Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,	125	145	Potential.
RA	Fishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,	146	154	Extracellular (Potential).
RA	Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,	155	175	Potential.
RA	Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,	176	176	Cytoplasmic (Potential).
RA	Sekhon L., Murray J., Sheet P., Cordes M., Abu-Threideh J.,	177	197	Potential.
RA	Stonking T., Kalicki J., Graves T., Harmon G., Edwards J.,	198	235	Extracellular (Potential).
RA	Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,	236	256	Potential.
RA	Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,	257	262	Cytoplasmic (Potential).
RA	Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,	263	283	Potential.
RA	Nelson J., Speth J., Ryan E., Andrews S., Geisel C., Layman D.,	284	315	Extracellular (Potential).
RA	Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,	316	336	Potential.
RA	Antonio B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,	337	354	Cytoplasmic (Potential).
RA	Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,	355	375	Potential.
RA	Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,	376	390	Extracellular (Potential).
RA	Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,			
RA	Chen E., Marra M.A., Martienssen R., McCombie W.R.;			
RT	"Sequence and analysis of chromosome 4 of the plant Arabidopsis			
RT	thaliana."			
RL	Nature 402:769-777(1999).			
CC	-!- FUNCTION: H(+)/sulfate cotransporter that may play a role in the			
CC	regulation of sulfate assimilation.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).			
CC	-!- TISSUE SPECIFICITY: Expressed only in leaves.			
CC	-!- SIMILARITY: Belongs to the SLC26A/Sulp transporter (TC 2.A.53)			
CC	family.			
CC	-!- SIMILARITY: Contains 1 STAS domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; AB004060; BAA20282.1; --			
DR	EMBL; AC002330; AAC78252.1; --			
DR	PIR; T01079; T01079.			
DR	GeneFam; 2358; 189.			
DR	InterPro; IPR002645; STAS.			
DR	InterPro; IPR001902; Sulph_transpt.			
DR	Pfam; PF01740; STAS; 1.			
DR	Pfam; PF00916; Sulfate_transp; 1.			
DR	TIGRFAMs; TIGR00815; sulP; 1.			
DR	PROSITE; PS01130; SLC26A; FALSE_NEG.			
DR	PROSITE; PS0801; STAS; 1.			
KW	Multigene family; Sulfate transport; Symport; Transmembrane;			
KW	Transport.			
FT	DOMAIN	1	76	Cytoplasmic (Potential).
FT	TRANSMEM	77	97	Potential.
FT	DOMAIN	98	99	Extracellular (Potential).
FT	TRANSMEM	100	120	Potential.
FT	DOMAIN	121	124	Cytoplasmic (Potential).
FT	TRANSMEM	125	145	Potential.
FT	DOMAIN	146	154	Extracellular (Potential).
FT	TRANSMEM	155	175	Potential.
FT	DOMAIN	176	176	Cytoplasmic (Potential).
FT	TRANSMEM	177	197	Potential.
FT	DOMAIN	198	235	Extracellular (Potential).
FT	TRANSMEM	236	256	Potential.
FT	DOMAIN	257	262	Cytoplasmic (Potential).
FT	TRANSMEM	263	283	Potential.
FT	DOMAIN	284	315	Extracellular (Potential).
FT	TRANSMEM	316	336	Potential.
FT	DOMAIN	337	354	Cytoplasmic (Potential).
FT	TRANSMEM	355	375	Potential.
FT	DOMAIN	376	390	Extracellular (Potential).

FT TRANSMEM 391 411 Potential.
FT TRANSMEM 412 432 Potential.
FT DOMAIN 433 447 Extracellular (Potential).
FT TRANSMEM 448 468 Potential.
FT DOMAIN 469 646 Cytoplasmic (Potential).
FT DOMAIN 504 627 STAS.
SQ SEQUENCE 646 AA; 71269 MW; B611E17A9B036E3 CRC64;

Query Match 66.6%; Score 2314; DB 1; Length 646;
Best Local Similarity 67.2%; Pred. No. 7.6e-144;
Matches 430; Conservative 104; Mismatches 102; Indels 4; Gaps 2;

Qy 42 HOVEVPPQPPFKSLKYSKLETFFDDPLQPKNK-PASKKFWLGLQFPPIFEWAPKYT 100
Db 10 HOVEIIPPQPPFLSKNTLNEILFADDPFRIRNESKTSKKIELGRHVPIILEWARGVS 69
Qy 101 FOFLKADLIAGITIASIAIPOGISYAKLANLPILGLYSYFIPPLIYAMGSSRDLAAGT 160
Db 70 LEYLKSDVISGIIIASIAIPOGISYAKLANLPILGLYSYSLVPLVYAINGSSRDLAAGT 129
Qy 161 VAGSLMGLMNAVDNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIIVDFLSHATI 220
Db 130 VAVASLLTAAMLGKEVNAVVPKLYLHLAFTATLFAAGLMQTCGLLRGLFVVEILSHAAI 189
Qy 221 IGFMGGAATVVCLOQLKSIILGLEHFTGADIISVMRSVFTQTHEWRESAVLGCVPFEL 280
Db 190 VGMGGAATVVCLOQLKGLLGLHFTHTSDIVTLRSIFSQSHMRWESVGLGCCFLIFL 249
Qy 281 LSTRYSKKRPREFWVSAMAPLTSVLGSLVFTFAEKHGVEVIGELKKGLNPPSLTNL 340
Db 250 LTTKYISKKPKLFWISMSPLVSVFTGTFILFLHDQFGIOFIELKKGINPPSIHL 309
Qy 341 VFVSPYMTTAVKTGIVVGIISLAEGIAVGRSPAMYKNYINIDGNKEMIAIGTMNVGSFTS 400
Db 310 VFTPPYVMLAKVGIITGVIALAEGIAVGRSPAMYKNYINIDGNKEMIAFGMMNLSFSS 369
Qy 401 CYLTGPFSSAVNAGCTAASNIIMSLAVMLTLFLPLPHYPLVLSAIIIVSAML 460
Db 370 CYLTGPFSSAVNAGCTALSNVMAVAVATLLFLPLPHYPLVLSAIIIAAML 429
Qy 461 GLDYEAAIHLFKVDKDFDFVVCMSAYIGVVFGSVEIGLVIATVISLRLFTARPTFV 520
Db 430 GLVDYEAAILHLKDKDFDFVVCMSAYIGVVFGSVEIGLVIATVISLRLFTARPTFV 489
Qy 521 LGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRERITRWIDEERIKAT 580
Db 490 MGNIQNSEIYRNIEHYPOAITRSSLLILHIDGPIYFANSTYLRDRIGRWIDEEDKLRTS 549
Qy 581 GETSLQVIIDMSAVGNIDTSGISMLSEVKKITERRELQVLVNPVSEVMMKLNKSKFON 640
Db 550 GDLSQVIIDMSAVGNIDTSGISMLSEELNKLILGRRELKIVIANPGAEVMMKLSKSTFIE 609
Qy 641 HLCKKWIYLTVEEAVGACNPLRASKTNPKKDETEGNNV 680
Db 610 SIGKERIYLTVAEAVACDFMLHTAKPDSVPPE---FNNV 646

RESULT 5
Q70BL1
ID Q70BL1 PRELIMINARY; PRT; 650 AA.
AC Q70BL1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Sulfate transporter.
GN Name=1ST3.2;
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.

RA Buchner P., Stuver E.E., Hawkesford M.J., de Kok L.J.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ601439; CAES3112.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR TIGRfam; TIGR00815; sulp; 1.
DR PROSITE; P50801; STAS; 1.
SQ SEQUENCE 650 AA; 71703 MW; 3880636AF83017DA CRC64;

Query Match 65.5%; Score 2274.5; DB 2; Length 650;
Best Local Similarity 65.4%; Pred. No. 3e-141;
Matches 428; Conservative 105; Mismatches 108; Indels 13; Gaps 4;

Qy 30 EYPLGMNFERVHVEVPPQPPFKSLKYSKLETFFDDPLQPKNK-PASKKFWLGLQF 88
Db 7 QYP-----QVETPPQPPFLSKNTLNEILFADDPFGKIRNESKTSKKIELGRH 56
Qy 89 FPIPEWAPKTYTQFLKADLIAGITIASIAIPOGISYAKLANLPILGLYSYFIPPLIYA 148
Db 57 VFPILEWAGYNLDYDKSDVISGIIIASIAIPOGISYAKLANLPILGLYSYFVPVMA 116
Qy 149 MNGSSRDLAAGTAVGSLMGLMNAVDNEDPKLYLHLAFTATLFAGVFOAALGLFRL 208
Db 117 IMGSSKDLAGIVAVASLLTAAMLGKEVSAVENPKLYLHLAFTSTFFAGLMQTCGLLR 176
Qy 209 GLIIVDFLSHATIIGFMGGAATVVCLOQLKSIILGLEHFTGADIISVMRSVFTQTHEW 268
Db 177 GFLVEITLSHAAIIGFMGGAATVVCLOQLKGLLGLHFTHTSDIVTLRSIFSQSPV 236
Qy 269 SAVLGCVPFELLSSTRYSKKRPREFWVSAMAPLTSVLGSLVFTFAEKHGVEVIGEL 328
Db 237 SGLLGCCLFLFELPTIKYISKKPKLFWISAMAPLSVIFGSLFVYFLHAQFHGIQIIGEL 296
Qy 329 KXGLNPPSIITNLVFSYPMYTTAVKTGIVVGIISLAEGIAVGRSPAMYKNYINIDGNKEMIA 388
Db 297 ERGINPPSIITNLVFSYPMYMLAKTGITGVIALAEGIAVGRSPAMYKNYINIDGNKEMIA 356
Qy 389 IGTNVVGSFTSCYLTGPFSSAVNAGCTAASNIIMSLAVMLTLFLPLPHYPL 448
Db 357 FGMNIFGSSCYLTGPFSSAVNAGCTAASNIIMSLAVMLTLFLPLPHYPL 416
Qy 449 VVLSAIIIVSAMLGLDYEAAIHLFKVDKDFDFVVCMSAYIGVVFGSVEIGLVIATVISL 508
Db 417 VVLSAIIIVSAMLGLDYEAAIHLFKVDKDFDFVVCMSAYIGVVFGSVEIGLVIATVISL 476
Qy 509 VLLFTARPTFVFLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRERITR 568
Db 477 LLLFVGRPKIYVNGNIQNTPEYNTIEQYQATLSSIIILHIDGPIYFANASYLRERIGR 536
Qy 569 WIDEERIKATGETSLQVIIDMSAVGNIDTSGISMLSEVKKITERRELQVLVNPVSE 628
Db 537 WIDEERIKATGETSLQVIIDMSAVGNIDTSGISMLSEVKKITERRELQVLVNPVSE 596
Qy 629 VMKLNKSKFQNLGKWIYLTVEEAVGACNPLRASK-TNPKKDETE-GWNNV 680
Db 597 LMKKLSKSKFPETIGKOWIHLTVAEAVSACDFVLTAKPDSPEKTSREPEFNNV 650

RESULT 6
Q7XE18
ID Q7XE18 PRELIMINARY; PRT; 658 AA.
AC Q7XE18;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to sulfate transporter (AT5G11).
GN ORFNames=OSJNBa0053G13.6;

```
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RL chromosome 10.";
RL Science 300:1566-1569 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wang R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017095; AAP53801.1; -.
DR Gramene; Q7XE18; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR003658; Antisig_antgst.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR ProDom; PD005210; Antisig_antgst; 1.
DR TIGFAMS; TIGR00815; sulp; 1.
DR PROSITE; P550801; STAS; 1.
SQ SEQUENCE 658 AA; 70652 MW; E7EF1315E2F22013 CRC64;

Query Match 59.2%; Score 2057; DB 2; Length 658;
Best Local Similarity 61.9%; Pred. No. 6.3e-127;
Matches 379; Conservative 104; Mismatches 129; Indels 0; Gaps 0;

QY 43 QVEVPPQPFKSLKYLKTEPPDPLRQFNKPKASKFMGLQFPFFIFWAPKYTFQ 102
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 20 RVVPVPPKPLRLTIGGNLMETFPDPPFRFVAVARESGRRALALRVVFPFLEWLP
QY 103 FLKADILAGITASTLAIPGISYAKLANLPILGLYSSFTPLIYAMGSSRDLAGTVA 162
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 80 ALMSDVAGVTITASLAVPOGISYAKUGLPPINGLYSSFPVLVAVMGSSRELAVGTTA 139
QY 163 VGSLLMGSMLSNAVDNEDPKLYHLHATFATFAGVFOALGFLRLGLVDFLSHATIG 222
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 140 VASLLFAATLGKAPGKPELYAALATFATFAGVLQAGLVRLGLFVLDLLSHAAIVG 199
QY 223 FMGGAATVCLQOLKSLGLGHEHTHGADIISVMSVFTQTHEWRWESAVLGCVFIFLLS 282
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 200 FMAGAAITVCLQOLKGLGLAHTTTSDVVAVVRSVVTQSHQWRWQSIYVVGCCFLIFL 259
QY 283 TRYFSKKRPRFFWVSAMAPLTSVILGSLVYFTTHAEKHGVEVIGELKGLNPPSLTNLYF 342
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 260 ARTISKRKPFNLWSAMAPLASVIAGSVLVLIHGDRHGIPVIGYKKGINPPSARDLL 319
QY 343 VSPYMTAVKTVGTVGIIISLAEGIAVGRSFAMFKYNNIDGNKEMIAIGTMVVGSTSCY 402
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 320 SPSHTTVALRTGIIITGIIAGLAEIAGRSFAMLSKSYNDGNKEMIAFGAMNIVGSCY 379
QY 403 LITGPPSRGANVYNAACKTAASNIIMSLAVMLTLFLPLFHYTPVLVLSAIVSAMGL 462
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 380 LTAGPFSRAAVNHNAGCKTPMSNAVAVMLTLQLFPLFHYTPVLVLSAIIISAMIGI 439
QY 463 IDYEAAILHFKVDKDFVFCMSAYIGVFGSVGEIGLVAIVLSVLRLPIARPRFVLG 522
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 440 IDYKAAVRLWKVDKIDFCVCGYLGWVFGDIOGLAIAVAGISILRILFIARPKTTVLG 499
QY 523 NIPNSVIYRNVEHYQNAKHPVGMILILEIDAPIYFANASYLRERITWIDEEERIKATGE 582
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
DB 500 KMPNSTNFRMDQYTVAKAVPGLLVLRIDSPISYFANSGLYLRERIMWIDHEEDRIKAEGL 559
QY 583 TSLQYVIIDMSAVGNTDTSIGSMLSEVKKITERRELQVLVNPVSVMKLNKSKFQNLH 642
DB :||||: : : : : : : : : : : : : : : : : : : : : : : : : :
-----
ID ST33 ARATH STANDARD; PRT; 631 AA.
AC Q9SXS2; O49307;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Probable sulfate transporter 3.3 (AST91)
GN Name=SULTR3;3; OrderedLocusNames=At1g23090; ORFNames=T26J12.13;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]_
RP SEQUENCE FROM N.A.
RA Takahashi H., Saito K.;
RT "Sulfate transporter AST91 from Arabidopsis thaliana.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ecv. Columbia;
RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pail G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820 (2000).
RN [3]
RP TISSUE SPECIFICITY.
RA Takahashi H., Watanabe A., Sasakura N., Aeanuma W., Nakamura M.,
RA Saito K.;
RT "Molecular characterization of the sulfate transporter gene family in
Arabidopsis thaliana.";
RL (In) Brunold C. (eds.);
RL Sulfur nutrition and sulfur assimilation in higher plants, pp.265-367,
Paul Haupt, Bern (2000).
CC -!- FUNCTION: H(+)/sulfate cotransporter that may play a role in the
regulation of sulfate assimilation (by similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Expressed only in leaves.
CC -!- SIMILARITY: Belongs to the SLC26A/Sulp transporter (TC 2.A.53)
family.
CC -!- SIMILARITY: Contains 1 STAS domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
enterprises requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
-----
CC      EMBL; AB023423; BAA75015.1; -.
DR      EMBL; AC002311; AAC00610.1; -.
DR      PIR; B86365; B86365.
DR      GeneFarm; 2362; 189.
DR      InterPro; IPR002645; STAS.
DR      InterPro; IPR001902; Sulph_transpt.
DR      Pfam; PF01740; STAS; 1.
DR      Pfam; PF00916; Sulfate_transp; 1.
DR      TIGRFAMs; TIGR00815; sulp; 1.
DR      PROSITE; PS01130; SLC26A; FALSE_NEG.
DR      PROSITE; PS0801; STAS; 1.
KW      Multigene family; Sulfate transport; Symport; Transmembrane;
KW      Transport.
FT      DOMAIN 1 69 Cytoplasmic (Potential).
FT      TRANSMEM 70 90 Potential.
FT      DOMAIN 91 92 Extracellular (Potential).
FT      TRANSMEM 93 113 Potential.
FT      DOMAIN 114 117 Cytoplasmic (Potential).
FT      TRANSMEM 118 138 Potential.
FT      DOMAIN 139 144 Extracellular (Potential).
FT      TRANSMEM 145 165 Potential.
FT      DOMAIN 166 171 Cytoplasmic (Potential).
FT      TRANSMEM 172 192 Potential.
FT      DOMAIN 193 223 Extracellular (Potential).
FT      TRANSMEM 224 244 Potential.
FT      DOMAIN 245 256 Cytoplasmic (Potential).
FT      TRANSMEM 257 277 Potential.
FT      DOMAIN 278 309 Extracellular (Potential).
FT      TRANSMEM 310 330 Potential.
FT      DOMAIN 331 347 Cytoplasmic (Potential).
FT      TRANSMEM 348 368 Potential.
FT      DOMAIN 369 384 Extracellular (Potential).
FT      TRANSMEM 385 405 Potential.
FT      DOMAIN 406 410 Cytoplasmic (Potential).
FT      TRANSMEM 411 431 Potential.
FT      DOMAIN 432 441 Extracellular (Potential).
FT      TRANSMEM 442 462 Potential.
FT      DOMAIN 463 631 Cytoplasmic (Potential).
FT      DOMAIN 497 621 STAS.
FT      CONFLICT 89 90 KL -> NV (in Ref. 1).
SQ      SEQUENCE 631 AA; 65070 MW; 21E48F67125E6200 CRC64;

Query Match          54.6%; Score 1898.5; DB 1; Length 631;
Best Local Similarity 57.1%; Pred. No. 1.6e-116;
Matches 353; Conservative 125; Mismatches 139; Indels 1; Gaps 1;

QY      41 VHOVEVPPQPFKSLKYLKETFPPDDPLRQFNKPKASKKFMGLGQFFPIPEWAPKYT 100
DB      3 VHKVAPPKSTVAKLTKLTKETFFDDPLRQFNKPKASKKFMGLGQFFPIPEWAPKYT 62
QY      101 FQFLKADLIAGITIASLAIPQGISYAKLANLPPIGLYSSFFIPPLIYAMGSSRDLAGVT 160
DB      63 FSLKSDVVSGLTISLAIPQGISYAKLANLPPIVGLYSSFFVPLVYAVLGSSRDLAGVP 122
QY      161 VAVGSLMLGMSLNAVDNEDPKLYLHLATATLPAQVFOAALGLFRGLIVDFLSHATI 220
DB      123 VSTASILGSMRLQVSPVDDPVLFLQLAFSSTFFAGLFOASILGILRLGFIIDFLSKATL 182
QY      221 IGFWGGAAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFQTQHEWESAVLGCVFIFL 280
DB      183 IGFWGGAAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFQTQHEWESAVLGCVFIFL 242
QY      281 LSTRYFSKKRPFWWSAMAPLTSVILGSLVLYFTTHAEKHGVRVIGELKKGLNPPSLTNL 340
DB      243 LSTRHLMSKKPKPLFWWSAGAPLISVIVSTLLVFFRAERHGISVIGKLPGLNPPSNMML 302
QY      341 VFVSPYMTTAVKTVGVVIGIISLAEGIAGVRSFAMRYKNYINDGNKEMIAIGTMNVGSGTS 400
DB      303 QFHGSHLALVAKTGLVTVIGISLATEGIAGVTRFAALKKNYHVDGNKEMIAIGLMNVGSGTS 362
QY      401 CYLTGPFSSAVNYAGCKTAASNIIMSLAVMLTLFLTPHLYPLVLSIIVSAML 460
DB      401 CYLTGPFSSAVNYAGCKTAASNIIMSLAVMLTLFLTPHLYPLVLSIIVSAML 460
-----
Db      363 CYTTGAFSRSVANNAGAKTAVSNIVMSVTVMVTLFLMPLPEYTPNVVLGAIIVTAVI 422
QY      461 GLIDYEAAJHLFKVDKDFVWCMSAYIGVWFGSVEIGLVIATIVISVLRLVLLFIARPTFV 520
DB      423 GLIDLPAACHIKIDKDFELVMLCAFPGVIFLSVQNGLAIVAGLSLFLKILMQVTRPKMVI 482
QY      521 LGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLREIRITRWIDEEERIKAT 580
DB      483 MGNIPGTDIYRDLHYHKEAQRIPGFLVLSIESPVNFANSNYLTERTSRWIECEESEAOE 542
QY      581 GETSQYVIIDMSAVGNIDTSGTSMLEEVKKITERRELQVLVNPVSEVMKLNKSKFQK 640
DB      543 KHSLSQLFLEMSAVSGVDVTNGVSPFKELKTKTAKKDIELVFNPLSEVVEKQLQRADEQK 602
QY      641 H-LGKKWYIYVEEAVGA 657
DB      603 EFMRPPEFLFTVAEAVAS 620

RESULT 8
Q9AT47 PRELIMINARY; PRT; 656 AA.
ID Q9AT47; AC Q9AT47;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sulfate transporter 2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=22983287; PubMed=129317983;
RA Howarth J R, Fourcroy P, Davidian J C., Smith F W., Hawkesford M J.;
RT "Cloning of two contrasting high-affinity sulfate transporters from
RT tomato induced by low sulfate and infection by the vascular pathogen
RT Verticillium dahliae.";
RL Planta 218:58-64(2003).
DR EMBL; AF347614; AAK27688.1; -.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0008271; P-sulfate porter activity; IEA.
DR GO; GO:0008272; P-sulfate transport; IEA.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS01130; SLC26A; 1.
DR PROSITE; PS0801; STAS; 1.
DR PROSITE; PS0801; STAS; 1.
SQ SEQUENCE 656 AA; 72556 MW; DE2A1DD2EF81C097 CRC64;

Query Match          54.1%; Score 1880.5; DB 2; Length 656;
Best Local Similarity 55.6%; Pred. No. 2.6e-115;
Matches 349; Conservative 119; Mismatches 159; Indels 1; Gaps 1;

QY      36 NNFERVHOVEVPPQPFKSLKYLKETFPPDDPLRQFNKPKASKKFMGLGQFFPIFEW 95
DB      25 NQTPYHKVGVPKQGVKFEKTKTVKETFADDPLRSFKDQPSRKLVLGLQAFILDW 84
QY      96 AFKTYTFQFLKADLIAGITIASLAIPQGISYAKLANLPPIGLYSSFFIPPLIYAMGSSRD 155
DB      85 GRSYNLRKFRGDLISGLTISLACIPQDIGYKLANLDPOGLYSSFPVPLVAFMGSSRD 144
QY      156 LAVGTAVVGSLLMGMSLNAVDNEDPKLYLHLATATLPAQVFOAALGLFRGLIVDFL 215
DB      145 IAIQVAVVSVLLGTLRLNEIDPSKHPAEYLAETATFAGITQATLGLRLGLIDFL 204
QY      216 SHATTIGFMGGAAATVVCLOQLKSIILGLEHFTHGADIISVMRSVF-QTQHEWESAVLGC 274
DB      205 SHAAVVGFMGGAAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFQHEWESAVLGC 264
```



```
Db 420 TLFLMPLFTTNNVILAAIILTAIVIGLIDYQAYKUMKVDKDFDFTCLCSFFGVLFVSV 479
Qy 495 EIGLVIAIVSLRVLFLFIARPTFVLGNIPNSVIYRNVEHYQNAKHVPQMLILEIDAPI 554
Db 480 PLGLATAVAVSVIKILLHVTRPNTSBEGNIPGQIQVSLGRYREASRIPGFLILAIIESPI 539
Qy 555 YFANASVLRERITRWIDEEERIKATGETSLQVVIIDMSAVGNIDTSGISMLSEVKKITE 614
Db 540 YFANSTYLODRILRLWAREREENRIKENNGTTLKCIILDMTAVSAIDTSGLVAEFLRRLE 599
Qy 615 RRELQVLNVPSEVMKLNKSFQNHGKKWLYLTVEEAVGACNFNLRS 665
Db 600 KQSLQVLNVPVGTVMKLRKSKIIIEALGSLGLYLTVEAVADLSSTWKAN 650

RESULT 10
Q629Y1 ID Q629Y1 PRELIMINARY; PRT; 656 AA.
AC Q629Y1; 2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative high affinity sulfate transporter.
GN Name=P0453D01.4;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004691; BAC98594.1; -.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS01130; SLC26A; 1.
DR PROSITE; PS0801; STAS; 1.
SQ SEQUENCE 656 AA; 72570 MW; 3FFB3A0D925B4E9 CRC64;

Query Match 53.5%; Score 1858; DB 2; Length 656;
Best Local Similarity 54.5%; Pred. No. 7.8e-114;
Matches 347; Conservative 121; Mismatches 165; Indels 4; Gaps 2;

Qy 34 GMNFERVHOVEVPPPPPKSLKYSLETFFDDPLRQPKNPKASKKFMGLQFFFPPIF 93
Db 23 GSSEERVYKRCPPQPKNFAREFRDTLRETFHNDPLRQYKQSGSAKFMALQFLPPIF 82

Qy 94 EWAPKTFQFLKADLIAGITIASLAIPOGISYAKLANLPILGLYSFPIPIYAMWGS 153
Db 83 EMGRCYNLRKFKGDLIAGLTIASLCIPQDYGSKLANLDAQYGLYSFVPPPIYAAWGS 142

Qy 154 RDLAVGTAVGSLMGLSMLSNVDPNEDPKLYLHLAFTATLFAGVFOALGLRGLIVD 213
Db 143 KDAIGVAVVSLIGSLQLNEVDVPKNKEEYLRLAFTATFFAGITQAAALGFLRGLIE 202

Qy 214 FLSHATIGPMGNAATVVCILQQLSKILGLEHFTHGADIISVMRSVFTQTHE-WRWESAVL 272
Db 203 FLSHAALVGMGNAATIALQQLKYLGIKSFYKTKTDIISVMRSVTSAHGNWQTVI 262

Qy 273 GCVPFIFPLSTRYFSKCRPFFVWSAMAPLTSVILGSLLYVTHAEKGVEIGELKGL 332
Db 263 GITFLAFLLLAKLYIGKKNRFFVPAIPATIPITSVILATLTFVITRADKQGVQVHNKGI 322

Qy 333 NPSSLNLVFPVSPYMTAVTGVIGVIGIISLAEGIAVGRSFAMYNKYNIDGNKEMIAIGTM 392
```

```
Db 323 NPSVHKHYFTGPFVAKGFKIGVISAMIGLTEAVAIGRTFAALKOYQLOGNKEMVALGTM 382
Qy 393 NVVGSFSTCYLTGTPFSRSVAVNAGCKTAASNIIMSLAVMLTLFLPLFHYTPTLVVLS 452
Db 383 NIAGSWTSCYIATGFSRSVAVNFMAGCQTPVSNINSVAVLLTLVITPLFKYTPNAILG 442
Qy 453 AIIVSAMLGLIDYEAIAHLFKVDKDFVVCMSAGYIGVFGSVBIGLVIAIVISVLRVLF 512
Db 443 SIIISAVIGLVYEAIVILIKVDMDFISCMGAFFGVFASVEIGLIIAVSISFAKILLQ 502
Qy 513 IARPTFVLGNIPNSVIYRNVEHYQNAKHVPQMLILEIDAPIYFANASVLRERITRWIDE 572
Db 503 VTRPRTVLLGNLPGTITTYRNTDOYPEARHIPGVIVRVDSAIYFSNSVYRRETLRWLTE 562
Qy 573 EERIKATGETSLQVVIIDMSAVGNIDTSGISMLSEVKKITERRELQVLNVPSEVMKK 632
Db 563 EEEKAKAEGSKINFLIEMSPVIDIDTSGIHALEDLYKNLKKRDIQLILANPGSIVMEK 622
Qy 633 LNKSKFQNHGKKWLYLTVEEAVGACNFNLRSKTNP 669
Db 623 LLSSKLNHEIGSNIFLTAVDAVRFT---RKSMQEP 656

RESULT 11
Q9AT12 ID Q9AT12 PRELIMINARY; PRT; 658 AA.
AC Q9AT12;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Sulfate transporter STL.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Hopkins L., Parmar S., Bouranis D.L., Howarth J.R., Hawkesford M.J.;
RT "Coordinated Expression of Sulfate Uptake and Components of the
RT Sulfate Assimilatory Pathway in Maize.";
RL Plant Biol. 6:408-414(2004).
DR EMBL; AF355602; AAK35215.1; -.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS0801; STAS; 1.
SQ SEQUENCE 658 AA; 72209 MW; 519DF5CB3DCC3412 CRC64;

Query Match 53.3%; Score 1850.5; DB 2; Length 658;
Best Local Similarity 56.1%; Pred. No. 2.5e-113;
Matches 345; Conservative 116; Mismatches 153; Indels 1; Gaps 1;

Qy 42 HQVEVPPPPPKSLKYSLETFFDDPLRQPKNPKASKKFMGLQFFFPPIFEMAKYTF 101
Db 33 YKGVPPPEKNLAEISDAVKTFDDPLRQYKQPSKKIWLGLQHIIPFVLDWSRHSYL 92

Qy 102 QFLKADLIAGITIASLAIPOGISYAKLANLPILGLYSFPIPIYAMWGSRLDLAGTV 161
Db 93 GFKPGDFIAGLTIASLCIPQDYGSKLANLPAEVLGYSFVPPPIYAMWGSRLDLAGTV 152

Qy 162 AVGSLMGLSMLSNVDPNEDPKLYLHLAFTATLFAGVFOALGLRGLIIVDFLSHATII 221
Db 153 AVVSLIGLGLQLNEVDVPKNKEEYLRLAFTATFFAGITQAAALGFLRGLIIFLSHAAIV 212

Qy 222 GFMGNAATVVCILQQLSKILGLEHFTHGADIISVMRSVFTQTHE-WRWESAVLCGVIFFL 280
```

```

Db 213 GFAGAAITIALQOLGFLGIANFTKKSIVSVKMSVGNVHGHWNQITLIGATLAF 272
Qy 281 LSTREYSKKPRFFWWSAMAPLTSVLGSLVYFTHAEKHGVIEGELKKGLNPPSLTNL 340
Db 273 LVAKYIGKRNKKLFWWSAIAPLTSVISTFTFVYITRADKHGVAIVKRNIRKNGINPPSASLT 332
Qy 341 VFVSPYMTTAVKGTGIVVGIISLAEGIAVGRSFAMYNKYNIDGNKEMIAIGTNNVGSFTS 400
Db 333 YFTGPIVATGFKIGIVAGMIGLTAIGTFAALKDYRIDGNKEMMALGTNNVGSLSIT 392
Qy 401 CYLTTGPFSSAVYNAGCKTAASNIIMSLAVMLTLLFLTPPHYTPLVVLSAIIIVSAML 460
Db 393 CYVATGFSRSVAVNMGCKTAVSNVMSIVVMVMTLLLTITPLPKYTPNALSSIIISAVL 452
Qy 461 GLIDYEAAHLFKVDKDFVWCMSAIVGVFGSVEIGLVIAIVISVLVLLFIARPTFV 520
Db 453 GLIDYESAVLIWKDKLDFLACGAPFGVIFSSVEYGLLIAVAISLAKILLQVTRPTVL 512
Qy 521 LGNIPNSVIRNVHYONAKHVPCMLILEIDAPIYFANASYLRERTWIDDEEBRIKAT 580
Db 513 LGNLPQTIYRNVEQYFDTAKVGVILVVDLSAIVFTNSYKVERILRWLRDEEBEQOQ 572
Qy 581 GETSLOVVIDMSAVGNIDTSGISMLEEVKKITRERELQLVLVNPVSEYMKLNKSKFQN 640
Db 573 KLTKEFLVLDLSPVIDIDTSGIHALAEELAKELKRIQLVLTNPVGPVAVIQKLSAKFTD 632
Qy 641 HLGKKWYLYLTVEEAV 655
Db 633 MIGEDNIFLTVGDVAV 647

RESULT 12
ST12_ARATH
ID ST12_ARATH STANDARD; PRT; 653 AA.
AC G9MAX3; Q9SH00;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Sulfate transporter 1.2.
GN Names:SULTR1.2; OrderedLocusNames=Atlg78000; ORFNames=F28K19.22;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND INDUCTION.
RC STRAIN=cv. Columbia;
RX MEDLINE=21836421; PubMed=11846879;
RA Yoshimoto N., Takahashi H., Smith F.W., Yamaya T., Saito K.;
RT "Two distinct high-affinity sulfate transporters with different
RL inducibilities mediate uptake of sulfate in Arabidopsis roots.";
RN Plant J. 29:465-473(2002).
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militschger J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,

```

```

RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RL thaliana";
RN Nature 408:816-820(2000).
RP [3]
RX FUNCTION, AND MUTANTS SEL1.
RX MEDLINE=21836422; PubMed=11846880;
RA Shibagaki N., Rose A., McDermott J.P., Fujiwara T., Hayashi H.,
RA Yoneyama T., Davies J.P.;
RT "Selenate-resistant mutants of Arabidopsis thaliana identify Sultr1.2,
RT a sulfate transporter required for efficient transport of sulfate into
RT roots.";
RL Plant J. 29:475-486(2002).
CC -!- FUNCTION: High-affinity H(+)/sulfate cotransporter that mediates
CC the uptake of the environmental sulfate by plant roots. Plays a
CC central role in the regulation of sulfate assimilation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in lateral root cap, root hairs,
CC epidermal and cortical cells of roots.
CC -!- INDUCTION: In roots by sulfate starvation.
CC -!- MISCELLANEOUS: Sell mutations in the gene lead to the resistance
CC of the plant to selenate, a toxic analog of sulfate.
CC -!- SIMILARITY: Belongs to the SLC26A/SulP transporter (TC 2.A.53)
CC family.
CC -!- SIMILARITY: Contains 1 STAS domain.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB042322; BAA95484.1; -.
CC EMBL; AC009243; AAF17685.1; ALT_SEQ.
CC PIR; D96809; D96809.
CC GeneFam; 2355; 189.
CC InterPro; IPR002645; STAS.
CC InterPro; IPR001902; Sulph_transpt.
CC Pfam; PF01740; STAS; 1.
CC Pfam; PF00916; Sulfate_transp; 1.
CC ProDom; PD006364; DUP52; 1.
CC TIGRFAMs; TIGR00815; sulp; 1.
CC PROSITE; PS01130; SLC26A; 1.
CC PROSITE; PS50801; STAS; 1.
CC Multigene family; Sulfate transport; Symport; Transmembrane;
CC Transport.
CC DOMAIN 1 91 Cytoplasmic (Potential).
CC TRANSMEM 92 112 Potential.
CC DOMAIN 113 116 Extracellular (Potential).
CC TRANSMEM 117 137 Potential.
CC DOMAIN 138 141 Cytoplasmic (Potential).
CC TRANSMEM 142 162 Potential.
CC DOMAIN 163 173 Extracellular (Potential).
CC TRANSMEM 174 194 Potential.
CC TRANSMEM 195 215 Potential.
CC DOMAIN 216 253 Extracellular (Potential).
CC TRANSMEM 254 274 Potential.
CC DOMAIN 275 301 Cytoplasmic (Potential).
CC TRANSMEM 281 301 Potential.
CC DOMAIN 302 339 Extracellular (Potential).
CC TRANSMEM 340 360 Potential.
CC DOMAIN 361 372 Cytoplasmic (Potential).
CC TRANSMEM 373 393 Potential.
CC DOMAIN 394 409 Extracellular (Potential).
CC TRANSMEM 410 430 Potential.
CC DOMAIN 431 438 Cytoplasmic (Potential).
CC TRANSMEM 439 459 Potential.
CC DOMAIN 460 466 Extracellular (Potential).
CC TRANSMEM 467 487 Potential.

```

FT	DOMAIN	488	653	Cytoplasmic (Potential).
FT	DOMAIN	522	645	STAS.
FT	MUTAGEN	511	511	I->T: In sell-8; reduces drastically sulfate transport activity.
SQ	SEQUENCE	653 AA; 71707 MW; B019CADD9AD9ED46	CRC64;	
Query Match				
Best Local Similarity 53.0%; Score 1841.5; DB 1; Length 653;				
Matches 343; Conservative 117; Mismatches 157; Indels 1; Gaps 1;				
Qy	42	HOVEVPPPPFRKSLKYSKTEFFPDPLQPKPKASKKMLGLQFPFFIFEWAPKYTF	101	
Db	28	HKVGIPPKQMPKDFMTTKETPFDDPDURDFKQPKSKQFMGLQSFVPFDMGRNVTIF	87	
Qy	102	QFLKADLIAGITIASLAIPQGISYAKLANLPILGLIYSFIPPLIYAMGSSRDIAVGTV	161	
Db	88	KKFRGDLISGLTIASLCIPQDIGYAKLANLPKYGLYSFVPLVYACWGSSRDIAIGPV	147	
Qy	162	AVGSLMGSMNSAVDPNEDPKLYHLATATLFAGVFOAALGRLGLIVDFLSHATII	221	
Db	148	AVVSLGLTLRAEIDPNTSPDEYLFALATATATFFAGITEAALGFRGLIDFLSHAUV	207	
Qy	222	GFNGGAATVVCLOAKSLGLEHFTHGADIIISVMSVFTQTHE-WRWESAVLGCVFIFPL	280	
Db	208	GFNGGAATIALQOLKGLGKFKTKTDIIISVLESVFKAAHHGNWQTILIGASFLTF	267	
Qy	281	LSTRYFSKKRPFVFWASMAPLTSVILGSLVYFTHAEKHGVEVIGELKGLNPPSLTNL	340	
Db	268	LTSKITGKSKKLFWPATAPLISVISTFVYITRADQGVQIVVKHLDQGINSSPHLI	327	
Qy	341	VFVSPYMTTAVKTGIVVIGISLAEGIAVGRSFAMYNKYNIDGNKEMIAIGTMNVVGSFTS	400	
Db	328	YFTGDNLAKGIRIGVAGVAGVALTEAVAIGKTFAMKDYQIDGNKEMVAGMNVVGSMS	387	
Qy	401	CYLTGTPFSRSVAVNAGCKTAASNIIMSLAVMLTLFLTPFHTPLVLSIAIVSAML	460	
Db	388	CYATGSFSRSVAVNFAGCOTAVSNIIIMSVLLTLFLTPFKYTPNAILAAIIIAVI	447	
Qy	461	GLIDYEAHLFKVDKDFVFCVMSAYIGVFGVSVEIGLVIATVSVLRLVFIARPRTFV	520	
Db	448	PLIDIQAAILFKVDKLDIACIGAFGVIFVSVIGILLIAVISFAKILLQVTRPTAV	507	
Qy	521	LGNIPNSVYRNVEHYONAKHVPGMILILEIDAPIYANASYLRERITRWIDEEBERIKAT	580	
Db	508	LGNIPRTSVYRNIIQYPEATVPGLTIRVDSAIYFNSVYRRIQRWLHEEBEKVAA	567	
Qy	581	GETSLQVYIIDSVAIGNIDTSGISMLSEVKKITERRELQVLVNPVSEVMKLNKSPON	640	
Db	568	SLPRIQFLIIEFSPVTDITDTSGHHALEDLYKSLQKRDITLILANPGPLVIGKHLHSHFAD	627	
Qy	641	HLGKKWYILTVBEAVGAC	658	
Db	628	MLGQDNIIYLTADAEAC	645	
RESULT 13				
ID	Q7X610	PRELIMINARY;	PRT;	629 AA.
AC	Q7X610;			
DT	01-OCT-2003	(TEMBLrel. 25, Created)		
DT	01-OCT-2003	(TEMBLrel. 25, Last sequence update)		
DT	05-JUL-2004	(TEMBLrel. 27, Last annotation update)		
DE	OSJNBa0060D06.5	protein (OSJNBa0059K02.23 proteoin).		
GN	Name=OSJNBa0060D06.5;	Synonyms=OSJNBa0059K02.23;		
OS	Oryza sativa	(japonica cultivar-group).		
OC	Eukaryota;	Viridiplantae; Streptophyta;		
OC	Spermatophyta;	Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartoideae;	Oryzeae; Oryza.		
OX	NCBI_TaxID=39947;			
RN	[1]			
SEQUENCE FROM N.A.				
RP	PubMed=12447439;	DOI=10.1038/nature01183;		
RA	Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,			

RP SEQUENCE FROM N.A.

RX PubMed=12447439; DOI=10.1038/nature01183;

RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,

100

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 31, 2005, 09:41:38 ; Search time 971 Seconds

(without alignments)
4145.648 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNTI.....NLRSKTNPKDETEGNWV 680

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10762049/runat_30082005_121244_5960/app.query.fasta_1.839
-DB=N Geneseq_16Dec04 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10762049 @CGN 1.1 470 @runat_30082005_121244_5960 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3475	100.0	2449	3	Aaz50489 Soybean s
2	2641.5	76.0	1977	3	Aac43310 Arabidops
3	2638.5	75.9	2112	3	Aac43093 Arabidops
4	2595.5	74.7	1977	6	Abz13697 Arabidops
5	2595.5	74.7	1977	8	Ada68514 Arabidops

6	2330.5	67.1	1974	8	ADA70002	Ada70002 Rice gene
7	2164	62.3	1981	3	Aaz50482	Aaz50482 Corn sulp
8	1902.5	54.7	2006	3	AAC42229	AAC42229 Arabidops
9	1901	54.7	2067	3	Aaz50488	Aaz50488 Soybean s
10	1898.5	54.6	1896	8	ADA68554	ADA68554 Arabidops
11	1844.5	53.1	2279	3	Aaz50481	Aaz50481 Corn sulp
12	1817	52.3	1940	8	ADA71315	ADA71315 Rice gene
13	1807.5	52.0	1371	6	ABQ82719	ABQ82719 Arabidops
14	1807.5	52.0	2000	8	ADA70110	ADA70110 Rice gene
15	1799	51.8	2311	3	Aaz50490	Aaz50490 Wheat sul
16	1777.5	51.2	1989	8	ADA70466	ADA70466 Rice gene
17	1683.5	48.4	1974	8	ADA70351	ADA70351 Rice gene
18	1675.5	48.2	1947	8	ADA69770	ADA69770 Rice gene
19	1644	47.3	1896	8	ADA69681	ADA69681 Rice gene
20	1513.5	43.6	1980	12	ADJ40066	ADJ40066 Plant cDN
21	1428	41.1	1983	8	ADA69804	ADA69804 Rice gene
22	1002	28.8	2058	6	ABZ12912	ABZ12912 Arabidops
23	1002	28.8	2058	8	ADA68540	ADA68540 Arabidops
24	963.5	27.7	2022	3	Aaz50491	Aaz50491 Wheat sul
25	731.5	21.1	780	3	Aaz50484	Aaz50484 Attichoke
26	706.5	20.3	2829	11	ADM56356	ADM56356 Frog SLC2
27	687.5	19.8	1791	11	ABD08637	ABD08637 Pseudomon
28	679.5	19.6	2612	11	ADM56273	ADM56273 Mouse SLC
29	679.5	19.6	2750	11	ADM56275	ADM56275 Mouse SLC
30	676.5	19.5	1767	11	ABD08580	ABD08580 Pseudomon
31	675.5	19.4	1240	3	Aaz50483	Aaz50483 Corn sulp
32	670	19.3	2279	11	ADM56358	ADM56358 Pig SLC36
33	668.5	19.2	2654	11	ADM56269	ADM56269 Human SLC
34	668	19.2	2589	13	ADQ87181	ADQ87181 Human tum
35	668	19.2	2748	11	ADM56271	ADM56271 Human SLC
36	665.5	19.2	2580	3	AA000611	AA000611 Human mem
37	664.5	19.1	2245	13	ACN43603	ACN43603 Human dia
38	662	19.1	2593	4	AAH75189	AAH75189 Nucleotid
39	662	19.1	12348	6	ABQ82327	ABQ82327 Human NOV
40	662	19.1	13935	5	AA574583	AA574583 DNA encod
41	659	19.0	2241	13	ACN43604	ACN43604 Human dia
42	658.5	18.9	2241	3	Aaz29466	Aaz29466 P. Chryso
43	657	18.9	2506	8	ABX12016	ABX12016 Transpore
44	651	18.7	2441	11	ADM57788	ADM57788 Murine pr
45	651	18.7	2441	12	ADE86099	ADE86099 cDNA enco

ALIGNMENTS

RESULT 1
AAZ50489
ID AAZ50489 standard; cDNA; 2449 BP.
XX
AC AAZ50489;
XX
DT 23-MAY-2000 (first entry)
XX
DE Soybean sulphate permease-2 cDNA clone.
XX
KW Sulphate Permease; sulphate assimilation protein; soybean; probe;
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
KW antibody; screen; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
CDS 2..2044
FT /*tag= a
FT /product= "Soybean sulphate permease-2"
FT /note= "Derived from clone sf11.pk0043.g10"
XX
PN WO200004154-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015810.
XX
PR 14-JUL-1998; 98US-0092833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PA Allen SM, Falco SC, Thorpe CJ;
 XX PI WPI; 2000-195025/17.
 XX DR P-PSDB; AAY44943.
 XX PT Nucleic acid fragments encoding sulfate assimilation proteins in plants
 PT and seeds useful as probes for isolating cDNAs and genes encoding
 PT homologous proteins, in producing transgenic plants.
 XX PS Claim 3; Page 52-53; 79pp; English.
 XX CC The present sequence is a cDNA clone encoding soybean sulphate permease,
 CC a sulphate assimilation protein. This sequence is obtained from sfl1
 CC library, clone sfl1.pk0043.g10, derived from soybean immature flower.
 CC This sequence is used as a probe to isolate other plant sulphate
 CC assimilation proteins, for genetic and physical mapping of related genes
 CC and as markers of traits linked to the gene. This is useful for plant
 CC breeding and to construct chimeric genes, used to create transgenic
 CC plants with altered levels of sulphate permease. The sulphate permease
 CC peptides are useful for producing antibodies, that are used to screen and
 CC isolate cDNA clones
 XX SQ Sequence 2449 BP; 719 A; 470 C; 543 G; 717 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	2449
Score:	3475.00	Matches:	680
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-762-049-18 (1-680) x AA250489 (1-2449)

XX	181	AspProLysLeuTyrlLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln	200
XX	542	GACCCAAAGCTTACCTCCACCTGGCTTTCACAGCTACATATTTGCTGGTGTTCAG	601
XX	201	AlaAlaLeuGlyLeuPheArgLeuGlyLeuValAspPheLeuSerHisAlaThrIle	220
XX	602	GCTGCCCTTGGGTCTGTTAGGTGGGGTGTGATCGTGGATTTCTGTGCATGCAACATA	661
XX	221	IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu	240
XX	662	ATAGGGTTTCATGGGAGGAGCAGCCAGGTGGTGTCTGCAGCAACTAAAAATCGATTCTT	721
XX	241	GlyLeuGluHisPheThrHisGlyAlaAlaSerIleIleSerValMetArgSerValPheThr	260
XX	722	GGCCTTGAGCATTTCAACCATGGAGCTGATATCATATCAGTCATGGCTCTGTTCACC	781
XX	261	GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu	280
XX	782	CAAACTCATGTAGTGAGGTGGAAAAGTGTGTGTAGGATGTCTCTCATTTCTTCCTC	841
XX	281	LeuSerThrArgTyrPheSerIlyLysArgProArgPhePheTrpValSerAlaMetAla	300
XX	842	CTTAGCACAAAGATATCTTCAGCAAAAACACCAAGGTTTTTTGGGTGTCACAATGGCG	901
XX	301	ProLeuThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHis	320
XX	902	CCATTGAGCTCGTTATATTGGGAAGTCTCTTGGTTTATTTCACCTCACGCCGAGAAGCAC	961
XX	321	GlyValGluValIleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsnLeu	340
XX	962	GGTGTGAAGTGATAGGAGAACTGAAGAAGGGTTTGAATCCACCATCATCAAAATCTG	1021
XX	341	ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle	360
XX	1022	GTATTTGTGTCGCTTACATGACTACAGCTGTCAAAACCTGGCATTTGCTGGCATCAT	1081
XX	361	SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle	380
XX	1082	TCACTTGGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTACAATATT	1141
XX	381	AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer	400
XX	1142	GATGGCAACAAAGAGATGATAGTATTGGGACCATGAAAGCTAGTTGGTTCTTTCACCTCT	1201
XX	401	CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys	420
XX	1202	TGCTACCTCACAAACAGGACCAATTTCCGGCTTCGGCTGTGAACATATAACCTGGATGCAAG	1261
XX	421	ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr	440
XX	1262	ACAGCAGCTTCCAAACATTAATATGTCACCTTGAGTAATGTTGACATTTGTTATTCCTGACA	1321
XX	441	ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValIleValSerAlaMetLeu	460
XX	1322	CCCTTGTTCATTAACATCCCTCGTGGTGTATCAGCTATTATCGTATCTCCAATGCTT	1381
XX	461	GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal	480
XX	1382	GGACTCATAGATTATGAAGCAGCCATCCATCTATTATTAAGTTGACAAATTTGACTTTGG	1441
XX	481	ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle	500
XX	1442	GTGTGCATGAGTGCATACATGTCGGTGTCTTTGGCAGTGTGTGAATTTGGCTTAGTCATA	1501
XX	501	AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal	520
XX	1502	GCTATTGTAATATCTGTACTTCGGGTACTTCTATTATTATGCAAGGCCAAGGACATTCGTT	1561
XX	521	LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys	540
XX	1562	TTGGGCAACATTCCAAATTCGTGATATACCGAAATTTGAGCACTATCAAAATGCATAA	1621
XX	541	HisValProGlyMetLeuLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer	560

Db	1622	CCTGGAATGCTAAATTCAGAGATTGATGCCCAATTTACTTTGCCAATGCCAGC	PR	14-MAY-1999;	99US-0134219P.
Qy	561	TyrLeuArgGluArgIleThrArgTrrPilleAspGluGluGluArgIleLysAlaThr	PR	14-MAY-1999;	99US-0134221P.
Db	1682	TAITTAAGACAAGGATCACAAAGTGGATTGATGAAGAAGAAAGAAATTAAGCTACA	PR	14-MAY-1999;	99US-0134370P.
Qy	581	GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr	PR	18-MAY-1999;	99US-0134768P.
Db	1742	GGGGAGACTAGTTTGGCAGTATGTTATAATTCATGATGCTCTGTGGAACATTTGATACA	PR	19-MAY-1999;	99US-0134941P.
Qy	601	SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeu	PR	20-MAY-1999;	99US-0135124P.
Db	1802	AGTGGATAAGTATGCTTGAAGAGGTGAAGAATTACAGAGAAGAGCTACAGCTT	PR	21-MAY-1999;	99US-0135353P.
Qy	621	ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn	PR	24-MAY-1999;	99US-0135629P.
Db	1862	GTTTTGGTCAATCCTGTGAAGTGAATGAAGAACTGAACAAATCGAAGTTCCAAAT	PR	25-MAY-1999;	99US-0136021P.
Qy	641	HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPhe	PR	27-MAY-1999;	99US-0136392P.
Db	1922	CATTTAGGGGAAGAAATGGATCTATCTGACTGTTGAAGAGGCCGTTGGAGCATGCAATTC	PR	28-MAY-1999;	99US-0137222P.
Qy	661	AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal	PR	03-JUN-1999;	99US-0137528P.
Db	1982	AACTACGTCAAGCAAAACCAACCCAAAGAAAGATGAACAGAGGTTGGAAACAATGTG	PR	04-JUN-1999;	99US-0137502P.
DE	RESULT 2				99US-0137724P.
XX	AAC43310				99US-0138094P.
XX	ID AAC43310 standard; DNA; 1977 BP.				99US-0138477P.
XX	AC AAC43310;				99US-0139119P.
XX	DT 17-OCT-2000 (first entry)				99US-0139452P.
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 38799.				99US-0139453P.
XX	Hybridisation assay; genetic mapping; gene expression control;				99US-0139454P.
XX	KW protein identification; signal transduction pathway; metabolic pathway;				99US-0139455P.
XX	KW promoter; termination sequence; ss.				99US-0139456P.
XX	OS Arabidopsis thaliana.				99US-0139457P.
XX	EP1033405-A2.				99US-0139458P.
XX	06-SEP-2000.				99US-0139459P.
XX	25-FEB-2000; 2000EP-00301439.				99US-0139460P.
PR	25-FEB-1999;	99US-0121825P.	PR	18-JUN-1999;	99US-0139461P.
PR	05-MAR-1999;	99US-0123180P.	PR	18-JUN-1999;	99US-0139462P.
PR	09-MAR-1999;	99US-0123548P.	PR	18-JUN-1999;	99US-0139463P.
PR	23-MAR-1999;	99US-0125788P.	PR	18-JUN-1999;	99US-0139464P.
PR	25-MAR-1999;	99US-0126264P.	PR	18-JUN-1999;	99US-0139465P.
PR	23-MAR-1999;	99US-0126785P.	PR	18-JUN-1999;	99US-0139466P.
PR	01-APR-1999;	99US-0127462P.	PR	18-JUN-1999;	99US-0139467P.
PR	06-APR-1999;	99US-0128234P.	PR	18-JUN-1999;	99US-0139468P.
PR	16-APR-1999;	99US-0128714P.	PR	18-JUN-1999;	99US-0139469P.
PR	19-APR-1999;	99US-0129845P.	PR	18-JUN-1999;	99US-0139470P.
PR	21-APR-1999;	99US-0130077P.	PR	18-JUN-1999;	99US-0139471P.
PR	23-APR-1999;	99US-0130510P.	PR	18-JUN-1999;	99US-0139472P.
PR	28-APR-1999;	99US-0130891P.	PR	18-JUN-1999;	99US-0139473P.
PR	30-APR-1999;	99US-0131449P.	PR	18-JUN-1999;	99US-0139474P.
PR	04-MAY-1999;	99US-0132048P.	PR	18-JUN-1999;	99US-0139475P.
PR	05-MAY-1999;	99US-0132484P.	PR	18-JUN-1999;	99US-0139476P.
PR	06-MAY-1999;	99US-0132485P.	PR	18-JUN-1999;	99US-0139477P.
PR	06-MAY-1999;	99US-0132486P.	PR	18-JUN-1999;	99US-0143542P.
PR	07-MAY-1999;	99US-0132487P.	PR	18-JUN-1999;	99US-0143543P.
PR	11-MAY-1999;	99US-0132863P.	PR	18-JUN-1999;	99US-0143544P.
PR	14-MAY-1999;	99US-0134256P.	PR	18-JUN-1999;	99US-0143545P.
PR	14-MAY-1999;	99US-0134218P.	PR	18-JUN-1999;	99US-0143546P.

```
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145913P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.

PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1.58e-262 Length: 1977
Score: 2641.50 Matches: 500
Percent Similarity: 88.16% Conservative: 81
Best Local Similarity: 75.87% Mismatches: 75
Query Match: 76.01% Indels: 3
DB: 3 Gaps: 3

US-10-762-049-18 (1-680) x AAC43310 (1-1977)

QY 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
Db 1 ATGGGCACGGAGGACTACACATTCCTCAAGGACGGAGGAGCTTCACCGCGCCACCAC 60
QY 44 ---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 61 ACGGTGGAGGCTCCCAACCTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120
QY 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 121 ACTCTGTTTCCAGACGACCTTTTAGACAAATTTAAGAACCCAAAATGCATCAAGAAAAT 180
QY 83 MetLeuGlyLeuGlnPhePhePheProLlePheGluTTPAlaProLysTyrThrPheGln 102
Db 181 GTGTTAGGCTCAAAATCTTCCTCCGATTTTCGAATGGGCACCGCTACCAATCTCAAG 240
QY 103 PheLeuLysAlaAspLeuAlaGlyLleThrLleAlaSerLeuAlaLleProGlnGly 122
Db 241 TTCTTCAAATCAGATCTCATCGCGGAATCACCATCGCTAGTCTCGCCATCCTCAGGGC 300
QY 123 IleSerTyrAlaLysLeuAlaAsnLeuProProLleLeuGlyLysTyrSerSerPheLle 142
Db 301 ATCAGTTAGCCCAAACTTTGCTAACTTGCCCCAAATTTTGGCCCTTTATTTCGAGTTTGA 360
QY 143 ProProLeuLleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 361 CCGCCATTGGTATACCGGTGCTAGGGAGTTCAAGGACTTAGCGGTGGGAACGGTTGCG 420
QY 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
Db 421 GTTGGCTCTCTGTTGACAGGTCCGATGCTGACAAAGAAAGTTGATGCTGAGAAAGATCCT 480
QY 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 481 AAGCTTTACCTTTACCTTTGCTTTTCCCGCCACTTTTTTCGCGGGGTTTCTCGAAGCTCT 540
QY 203 LeuGlyLeuPheArgLeuGlyLeuLleValAspPheLeuSerHisAlaThrLleLleGly 222
Db 541 CTTGGAATTTTCAGGTTAGGGTTTCATAGTGGACTTTCTATCCGATGCCAACGATAGTAGA 600
QY 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerLleLeuGlyLeu 242
Db 601 TTCATGGGAGGAGCAGCGAGCGGTGGTGTGCTGTCAACAGCTTAAAGGTATTTTCGACTT 660
QY 243 GluHisPheThrHisGlyAlaAspLleLleSerValMetArgSerValPheThrGlnThr 262
Db 661 AAACATTTTCACAGACTCTACCGATGTATCTCTGTGTCGCTCCGTTCTCCCAAACT 720
QY 263 HisGluTTPArgTTPGluSerAlaValLeuGlyCysValPhePheLlePhePheLeuLeuSer 282
Db 721 CACGAGTGGAGATGGGAAAGTGGCGCTTTTGGATGGTGGTTTCTATTCTTCTCTCTCC 780
```

283 ThrArgTyrPheSerLysLysArgProArgPhePheTyrValSerAlaMetAlaProLeu 302
Db ACCAGATATTTTCAGCATCAAGAAACCAAAATCTTTTGGGTGGCGGATGGCTCTCTTTG 840
303 ThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVal 322
Db ACCTCAGTGATCTTGGAGTCTCTTGGTTTACTTCACTCACGCTGAGAGACATGCTGTT 900
323 GluValIleGlyLysLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPhe 342
Db CAAGTGTATAGGACCTGGAAGAGGGTTGAATCCACTCTCCGGTCTTGATCTCATCTTT 960
343 ValSerProTyrMetThrThrAlaValLysThrGlyValValGlyIleIleSerLeu 362
Db ACTTCCCTTACATGTCACAGCTGTCAAACTGGGCTCATCTGCGCATCTTGCTCTC 1020
363 AlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly 382
Db GCTGAAGGAGTACAGTGGGAGGAGTTTGGCATGTTCAAGAACTACAACATAGACGGG 1080
383 AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402
Db AACAAAGAGATGATAGCGTTTGAATGATGAACATCGTTGCTTCCATCTGCTTAC 1140
403 LeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla 422
Db CTCACAACCGGACCATTTTCAAGTCGGCAGTGAACCTACAACGCGGGTTGCAAGACCGCA 1200
423 AlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuProLeu 442
Db ATGTCCACACATAGTATGGGATGGGTATGTTTACACTCTCTCTCTCACACCGCTT 1260
443 PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu 462
Db TTTCACTACACACCACTCGTCGCTCTCTGCCATCATCATATCCGAATGCTCGGACTC 1320
463 IleAspTyrClnuAlaIleHisLeuPheLysValAspLysPheAspValValCys 482
Db ATTGACTATCAAGCTCCATCTCTGGAAGTTGACAAAGTTCCACTTCCTCGTCTGC 1380
483 MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle 502
Db ATGAGCGCTACGTTGGGTCTGTTATTCGGCAGTGTAGAGATTGGATCTCGTAGCGGTG 1440
503 ValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGly 522
Db CGCATATCTATAGCAGGTGTTGCTGTTGTGTCGAGGCCAAAACTGCGGTGAAGGGA 1500
523 AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVal 542
Db AACATACCAACAGCATGATCTATAGGAACACTGAGCAGTACCCATCATCAAGAACCGTT 1560
543 ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu 562
Db CTGGTATCTCATCTTGGAGATTGATGCTCCCATCTACTTGTCTAATGCCAGTTACTTG 1620
563 ArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThrGlyGlu 582
Db CGTGAAGAATCATAAAGTGGATTGATGAAGAGAGAGAGAGTTAAACAATCAGGAGAG 1680
583 ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly 602
Db AGCAGCTTCAATATATATTACTCGATATGTCAGTGTGTTGTTAATATCGACACACCGGT 1740
603 IleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLeu 622
Db ATTAGCATGTGTGGAAATTAAGAAGTCAATGACAGGAGACCGTTAAGATTGGTATTG 1800
623 ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641
Db TCAAAATCCAAAGAGAGGTCGTAAGAAATTAACCAAGATCCAAATTCATCGGTGATCAT 1860
642 LeuGlyLysLysTrpIleTyrLeuThrValGluAlaValGlyAlaCysAsnPheAsn 661

1861 TTGGGCAAGAGTGGATGTCTTAAACGGTAGGAGAACGAGTGGAGCTTGTAGCTACATG 1920
662 LeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db CTTTCACACGTTTAAACACCGAACCG---GCCTCCAAAACGAGCCTTGGAAACACGTA 1974
RESULT 3
AAC43093
ID AAC43093 standard; DNA; 2112 BP.
XX AAC43093;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 38002.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 01-APR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158365P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. NO.: 3.57e-262 Length: 2112
Score: 2638.50 Matches: 499
Percent Similarity: 88.16% Conservative: 82
Best Local Similarity: 75.72% Mismatches: 75
Query Match: 75.93% Indels: 3
DB: 3 Gaps: 3

US-10-762-049-18 (1-680) x AAC43093 (1-2112)

XX WO200216655-A2.
XX
XX
XX PD
XX
XX PF 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US026685.
XX
XX 24-AUG-2000; 2000US-0227866P.
XX PR 26-JAN-2001; 2001US-0264647P.
XX PR 22-JUN-2001; 2001US-0300111P.
XX
XX (SCRI) SCRIPPS RES INST.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
XX
XX WPI; 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed and
XX PT producing plants with increased tolerance to these abiotic stresses.
XX
XX PS Claim 144; SEQ ID NO 1502; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
XX CC cell has been exposed, comprising: (a) contacting nucleic acid
XX CC representative of expressed polynucleotides in the plant cell with an
XX CC array or probes representative of the plant cell genome; and (b)
XX CC detecting a profile of expressed polynucleotides in the plant cell
XX CC characteristic of a stress response. The method is useful in the
XX CC production of transgenic plants, cells and seeds and in producing plants
XX CC with increased tolerance to abiotic stress. The present sequence is that
XX CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX CC in methods of the invention. Note: The sequence data for this patent is
XX CC not represented in the printed specification but is based on sequence
XX CC information supplied to Derwent by the European Patent Office
XX
XX SQ Sequence 1977 BP; 513 A; 458 C; 471 G; 535 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9,06e-258 Length: 1977
Score: 2595.50 Matches: 498
Percent Similarity: 87.73% Conservative: 81
Best Local Similarity: 75.45% Mismatches: 77
Query Match: 74.69% Indels: 5
DB: 6 Gaps: 3

US-10-762-049-18 (1-680) x ABZ13697 (1-1977)

QY 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
DB 1 ATGGGCACGGAGGACTACACATTCCTCAAGGAGCGGAGGAGCTTCACCGCGCCACAC 60

QY 44 ---ValGluValProProGlnProPhePheLeuSerLeuLysTyrSerLeuGlu 62
DB 61 ACGGTGGAGGCTCCACACCTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120

QY 63 ThrPhePheProAspProLeuAArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
DB 121 ACTCTGTTCCAGACGACCTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 180

QY 83 MetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLysTyrThrPheGln 102
DB 181 GTGTTAGGCTCAAAATCTCTCCCGATTTTCGAATGGGCACACGCTACAATCTCAAG 240

QY 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
DB 241 TTCTTCAAAATCAGATCTCATCGCGGAATCACCATCGCTAGTCTCGCCATCCCTCAGGC 300

QY 123 IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerPheIle 142
DB 301 ATCAGTTACGCCAAACTTGTACTACTTGGCCCCCAATTCCTGGCCCTTATTCGAGTTTGA 360

QY 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162

DB 361 CCGCCATTGGTATACGCGGTGCTAGGGAGTTCAAGGGACTTAGCGGTGGGAACGGTGGC 420
QY ValGlySerLeuLeuMetClySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
DB 421 GTTGGCTCTCTGTTGACAGGTGCGATGCTGACAAAGAAAGTGTGATGCTGAGAAAGATCT 480
QY LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
DB 481 AAGCTTTACCTTCACCTTGCTTTCACCGCCACTTTTTCGCCGGCGTCTCGAAGCCTCT 540
QY LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleG 222
DB 541 CTTGGAATTTTCAGGGTTAGGGTTTCATAGTGACTTCTATCGCATGCAACGATAGTAGG 600
QY 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeuGlyLe 242
DB 601 ATTTCATGGGAGGAGCAGCGGCTGAGTCTGCAACAGCTTAAGGGTATTTTCGGACT 660
QY 242 uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnTh 262
DB 661 TAAACATTTTCACAGACTCTACCGATGTTATCTCTGTGATCGCTTCGGTTTCTCCAAAC 720
QY 262 rHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
DB 721 TCACGAGTGGAGATGGAAAGTGGCGTCTTGATGTGGTGTGTTTCTATTCTTCTCTCTC 780
QY 282 rThrArgTyrPheSerLysLysArgProArgPheThrValSerAlaMetAlaProLe 302
DB 781 CACCGAGATATTGAGCATCAAGAAACCAAAATCTTTTGGGTGGCGGCGATGGCTCTCTT 840
QY 302 uThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVa 322
DB 841 GACCTCAGTGATCTTGGAGTCTCTTGGTTACTTCACTCAGCTGAGAGACATGGTCT 900
QY 322 lGluValIleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342
DB 901 TCAAGTGATAGGGACCTGAAGAAAGGGTTGAATCCACTCTCCGGTCTGTGATCTCATCTT 960
QY 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValGlyIleIleSerLe 362
DB 961 TACTTCCCCCTTACATGTCCACAGCTGTCAAAACTGCGCTCATCATCTGGCATCATTTGCTCT 1020
QY 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspG 382
DB 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTCGCGATGTTCAAGAACTACAACATAGACGG 1079
QY 382 yAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCyst 402
DB 1080 GAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCTCTTCACATCTTGT 1139
QY 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422
DB 1140 CTTCAACACCGGACCATTTTCAAGTGGCAGTGAATCAACACGGGGTTGCAAGCCGC 1199
QY 422 aAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLe 442
DB 1200 AATGTCCAACATAGTATGATGGCGATTGCGGTTATGTTTACACTCTCTTCTCACACCGCT 1259
QY 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462
DB 1260 TTTTCACTACACACCATCGTCTCTCTGCCATCATCATATCGCAATGCTCGGACT 1319
QY 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482
DB 1320 CATTCATCATCAAGCTGCCATCATCTCTGAAAGTTGCAAGTTGCAAGTTCTCTGCTG 1379
QY 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla 502
DB 1380 CATGAGCGCTACGTTGGGTGCTATTCGGCAGTGTAGAGATTGGACTCGTCTGTCGCTG 1439
QY 502 eValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuG 522
;

Db 1440 GCGATATCTATACGAGGTGTTGCTGTTGTCGAGGCCAAAACTCGGTGAAGG 1499
Qy 522 yAsnIleProAsnSerValIleTyArgAsnValGluHieTyGlnAsnAlaLysHisVa 542
Db 1500 AAACATACCAACAGCATGATCTATAGGAACACTGACGAGTACCCTCATCAAGAACGT 1559
Qy 542 lProGlyMetLeuIleLeuGluIleAspAlaProIleTyPheAlaAsnAlaSerTyLe 562
Db 1560 TCCTGGTATCTCATCTGAGATGATGCTCCCATCTCTTGTCTAATGCCAGTTACTT 1619
Qy 562 uArgGluArgIleThrArgTrpIleAspGluGluGluGluArgIleLysAlaThrGlyGl 582
Db 1620 GCGTGAAGAAATCATAGGTGGATTGATGAAGAGAGAGAGAGTAAACATCAGGAGA 1679
Qy 582 uThrSerLeuGlnTyValIlelleAspMetSerAlaValGlyAsnIleAspThrSerGl 602
Db 1680 GAGCAGCTTCAATATATATATCTCGATATGTCAGCTGTGTGTAATATCGACCAAGCG 1739
Qy 602 yIleSerMetLeuGluGluValIleValIleThrGluArgArgGluLeuGlnLeuValle 622
Db 1740 TATTAGCATGATGTTGGAATTAAGAAAGTCATTGACAGAGAGCGGTAAAGTTGGTATT 1799
Qy 622 uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe--GlnAsnHi 641
Db 1800 GTCAATCCAAAGGAGAGGTGCGTGAAGAAATTAACAGATCCAAATTCATCGTGATCA 1859
Qy 641 sLeuGlyLysLeuTrpIleTyLeuThrValGluGluAlaValGlyAlaCysAsnPheAs 661
Db 1860 TTTGGGCAAGAGTGGATGTTCTTAACGGTAGGAGAGCAGTGGAGGCTTGTAGCTACAT 1919
Qy 661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 1920 GCTTCACAGCTTTAAACCAACCG--GCCTCAAAACAGAGCGCTTGGAAACAACGTA 1974
RESULT 5
ADA68514
ID ADA68514 standard; DNA; 1977 BP.
XX AC ADA68514;
XX DT 20-NOV-2003 (first entry)
XX DE Arabidopsis thaliana gene, SEQ ID 589.
XX KW Plant; bacterial infection; fungal infection; viral infection; ds.
XX OS Arabidopsis thaliana.
XX PN W02003000898-A1.
XX PD 03-JAN-2003.
XX PF 22-JUN-2001; 2001WO-IB001105.
XX PR 22-JUN-2001; 2001WO-IB001105.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX DR WPI; 2003-175290/17.
XX PT Identifying at least one gene involved in plant resistance or response to
XX PT pathogenic infection for conferring resistance or tolerance to a plant to
XX PT bacterial, fungal or viral infection by determining or detecting plant
XX PT gene expression.
XX PS Claim 6; SEQ ID NO 589; 899pp; English.
XX CC The present invention relates to a method (M1) for identifying genes
XX CC involved in plant resistance or response to pathogenic infection. M1
XX CC comprises identifying a gene whose expression is significantly altered in

CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 1977 BP; 513 A; 458 C; 471 G; 535 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 9,06e-258 Length: 1977
Score: 2595.50 Matches: 498
Percent Similarity: 87.73% Conservative: 81
Best Local Similarity: 75.45% Mismatches: 77
Query Match: 74.69% Indels: 5
DB: 8 Gaps: 3
US-10-762-049-18 (1-680) x ADA68514 (1-1977)
Qy 24 MetGlySerValAspTyArgTrpIleProLeuGlyMetAsnAsnPheGluArgValHisGln 43
Db 1 ATGGGCAAGGAGGAGTACACATTCCCTCAAGGAGCGGAGGAGCTTCCAGCGCCGCCAC 60
Qy 44 ---ValGluValProProGlnProPhePheLysSerLeuLysTySerLeuLysGlu 62
Db 61 ACGGTGGAGGCTCCACAACTCAACCGTTCTTGAAGTCATCTTCAGTACTCAGTGAAGAA 120
Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 121 ACTCTGTTCCAGACGACCCCTTTAGACAATTTAAGAACCAAAATGCATCAGAAATTT 180
Qy 83 MetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyThrPheGln 102
Db 181 GTGTAGGCTCAAAATACTTCTCCCGATTTCGNAATGGCCACCGCTACAAATCTCAAG 240
Qy 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 241 TTTCTCAAAATCAGATCTCATCGCGGAATCACCATCGCTAGTCTCGCCATCCCTCAGGC 300
Qy 123 IleSerTyAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTySerSerPheIle 142
Db 301 ATCAGTTACGCCAAACTTGTAACTTGGCCCCCAATTTCTGGCTTTATTCGAGTTTGT 360
Qy 143 ProProLeuIleTyArgAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 361 CGCCATTGGTATACGCGTGTAGGAGTCAAGGAGCTTCAAGGAGCTTAGCGGTGGACGGTTCG 420
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
Db 421 GTTGGTCTCTGTGACAGTGGATGCTGAGCAAGAAAGATTGATGCTGAGAAAGATCTCT 480
Qy 183 LysLeuTyLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 481 AAGCTTTACCTTCACCTTCTTCCCGCCACTTTTTCGCGCGCGTTCCTCGAAGCTCT 540
Qy 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGl 222
Db 541 CTTGGAATTTTCAGGGTTAGGGTTTCATAGTGAGCTTTCTATCGCATGCAACGATAGTAG 600
Qy 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242
Db 601 ATTATGGGAGGAGCAGCGGTGTGTGAGTCTGCAACAGCTTAAAGGTATTTTCGAGCT 660
Qy 242 uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnTh 262
Db 661 TAAACATTTCCAGACTCTACCGATGTTATCTCTGTGATGCGTTCCTGTTCTCCCAAC 720
Qy 262 rHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
Db 721 TCACGAGTGGAGATGGAAAGTGGCGTCTTGGATGTGGTCTTCTATTCTTCTTCTC 780
Qy 282 rThrArgTyPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProle 302

[illegible]

Db 1860 TTTGGGCAAGAGTGGATGTCTTTAAACGGTAGGAGAAGCAGTCGGAGCGCTTGTTAGCTACAT 1919

Qy 661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
||| ||||| ||| :||| ||||| |||||

Dd 1920 GCTTCACAGTGTAAACCGAACCG---GCCTCCAAAACGAGCCTTGGACACGTA 1974

RESULT 6

ADA70002

ID ADA70002 standard; DNA; 1974 BP.

XX AC ADA70002;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 3325.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice;
gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175290/17.

XX PT Identifying at least one gene involved in plant resistance or response to
pathogenic infection for conferring resistance or tolerance to a plant to
bacterial, fungal or viral infection by determining or detecting plant
gene expression.

XX PS Claim 6; SEQ ID NO 3325; 899pp; English.

XX CC The present invention relates to a method (M1) for identifying genes
involved in plant resistance or response to pathogenic infection. M1
comprises identifying a gene whose expression is significantly altered in
the incompatible interaction of plant gene expression relative to
expression of the gene in an uninfected plant, in a mutant plant that
does not express a gene associated with response to pathogenic infection,
or in a corresponding incompatible or compatible interaction. (M1) is
useful for conferring resistance to resistance or tolerance to a plant to
bacterial, fungal or viral infection. The present sequence was used to
illustrate the invention.

SQ Sequence 1974 BP; 335 A; 602 C; 657 G; 380 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2_3le-230	Length:	1974
Score:	2330.50	Matches:	441
Percent Similarity:	83.84%	Conservative:	83
Best Local Similarity:	70.56%	Mismatches:	100
Query Match:	67.06%	Indels:	1
DB:	8	Gaps:	1

US-10-762-049-18 (1-680) x ADA70002 (1-1974)

Qy 43 GlnValGluValProProGlnProPhePheLysSerLysLysTyrrSerLeuLysGlu 62
:::||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||

Dd 52 AGGGTCGGATGCCGGCGGAAGCCGTTCTCGAGACGCTGGGGGGGAACAATGAAGGAG 111
:::||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||

Qy 63 ThrPhePheProAspAspProLeuArgGlnPhe---LysAsnLysProAlaSerLysLys 81
:::||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||

Dd 112 ACAATTCTCGCGACACCACCGGTTTCAGGGTGTTGTCGGCGGAGCGGGGTGCGGGCGCGC 171
:::||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||


```
QY 82 PheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPhe 101
Db |||:|||||
172 GCGCGCGCGCGCTCCGGTAGCTGCCGTTTCATGGAGTGGCGCGCTGCATCACCCCTC 231
QY 102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db |||:|||||
232 GGCACCCCTCAAGTCCGACCTCATCGCGGCATCACCATTCGCAGCCTCGCCATCCCCAG 291
QY 122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPhe 141
Db |||:|||||
292 GGCATCAGCTACGCAAGCTCGCCAACTCCCTCCCTCCCTCGCTTCATTCGAGCTTC 351
QY 142 IleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db |||:|||||
352 GTGCGCGCGCTGGTACGCGATGATGGGGAGCTCGAGGGACCTCGCGGTGGGAGCGGTG 411
QY 162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db |||:|||||
412 GCGTGGCGTCCGTGCTGATCGGTTCGATGTCGAGCGAGGAGTGTGCGCGCGGAGGAC 471
QY 182 ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db |||:|||||
472 CCGCGCGCTGATCTGCACTGCGCTCAGCGCCACCTTCTTGGCGCGGTTCACAGCG 531
QY 202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db |||:|||||
532 CTGCTGGGGGCTTGGAGCTGGGGTTCATGTCGACTTCCTGTCGACCGCACCATCGTC 591
QY 222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db |||:|||||
592 GGGTTCATGGCGCGCGCCACCGCTGGTGTCTGCGAGCGAGCTCAAGGGCATGTTCCGC 651
QY 242 LeuGluHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThrGln 261
Db |||:|||||
652 CTCGACCACCTCACCAACCGCCACCGACTCGTCTCCGTCATGAGCTCCGTTCTCCCAA 711
QY 262 ThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeuLeu 281
Db |||:|||||
712 ACCACCTCTGGCGATGGGAGAGCTGCTCATGGGCTGCGGCTTCCTCTTCTTCCTCTC 771
QY 282 SerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaPro 301
Db |||:|||||
772 ATCACCGCTCTTCACAGAGAGGAGGCAAGGTTCTTCGGGTATCTGAGCTGCGCCCA 831
QY 302 LeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGly 321
Db |||:|||||
832 TTGGCGTCTGTTCATCATCGGAGCTCTCTGTGTACCTCACACATGCTGAAACCATGGC 891
QY 322 ValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuVal 341
Db |||:|||||
892 ATTCAAGTGTATTGTTTACCTGAAGAAGGCGCTGAACCCCGCTGCGCACCAAGCTGAAC 951
QY 342 PheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSer 361
Db |||:|||||
952 TTCTCTGCGCGTACATGATGCTGCGCTCGAGACCGGATCATCATCTGCGGTATTGCC 1011
QY 362 LeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAsp 381
Db |||:|||||
1012 CTCGCTGAGGGGATTGCGGTAGGAAGGAGCTTTGCAATGTTCAAGAATTACCATCATCGAC 1071
QY 382 GlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCys 401
Db |||:|||||
1072 GGAACAAGGAGATGATCGCGTTTGGAAACGATGAACATTGTGGGATCATCTCACCTCTTGC 1131
QY 402 TyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThr 421
Db |||:|||||
1132 TACCTTACCACCGGTCTGTTCTCGAGGTGGCGGTCAACTACACCGCGGCTGCAAGACG 1191
QY 422 AlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrPro 441
Db |||:|||||
1192 GCGATGTGGAACGTGATCATGTCGGTGGCGGTGATGATCACGCTGCTGTTCTTGACGCG 1251
```

```
QY 442 LeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGly 461
Db |||:|||||
1252 CTGTTCCACTACACCGCGCTGGTGGTCTGTCGCGCATCATGTCGCGCATGCTCGGG 1311
QY 462 LeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValVal 481
Db |||:|||||
1312 CTGATCGACTACCGCGCGCGCTCCACCTGTGGCAGGTGCACAAGGTGGACTTCTCGGTC 1371
QY 482 CysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla 501
Db |||:|||||
1372 TGCTCGCGCTTACCTCGCTCGCTCGCTGCTTCCGCGAGGTGAGATCGGCTCGTCTGCC 1431
QY 502 IleValIleSerValLeuLeuPheIleAlaArgProArgThrPheValLeu 521
Db |||:|||||
1432 GTCGGATTCCTCATCTCCGCTGCTGCTGTCGCGCGCGCGAGGACGAGCTGCTC 1491
QY 522 GlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHis 541
Db |||:|||||
1492 GGGAAACATCCCAACACGATGATCTACCGCGCGATGGACCATACACCGCGCGCAGG 1551
QY 542 ValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyr 561
Db |||:|||||
1552 GTCCCGCGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1611
QY 562 LeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGly 581
Db |||:|||||
1612 CTGCGTGGAGGATCGCGCTGGATCGACGAGGAGGACCATGTCGCAAGGAGAGGCG 1671
QY 582 GluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSer 601
Db |||:|||||
1672 GAGATGGCATTCAGTACGTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1731
QY 602 GlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeuVal 621
Db |||:|||||
1732 GGGACGAGCATGCTGGATGAATCAGAGAGACCTCGACAGGAGGCGCTTCAGATTGTG 1791
QY 622 LeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHis 641
Db |||:|||||
1792 TTGCGAACCCTGGGAGCGAGATCATGAAGAAGTTGACAGCTCCCAAGGTGCTTGAGCG 1851
QY 642 LeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661
Db |||:|||||
1852 ATCGGCATGATGATGATCTTCCCACTGGTGGCGAGCGGTGGCGGAGTCCGACTTCGT 1911
QY 662 LeuArgAlaSerLys 666
Db |||:|||||
1912 ATGCACTCGCAGAAG 1926
RESULT 7
AAZ50482 standard; cDNA; 1981 BP.
XX
AC AAZ50482;
XX
DT 23-MAY-2000 (first entry)
XX
DE Corn sulphate permease-2 cDNA clone.
XX
KW Sulphate Permease; sulphate assimilation protein; corn; contig; probe;
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
KW antibody; screen; ss.
XX
OS Zea mays.
XX
FH Key
CDS 2..1741
FT /tag= a
FT /product= "Corn sulphate permease-2"
FT /note= "Derived from contig composed of clones
FT crln.pk0015.a2, p0006.cbyvs25rb, p0072.comhc25r,
FT p0091.cmard29r and p0092.chwat43r"
XX
PN WO200004154-A2.
```

```
XX 27-JAN-2000.
XX 13-JUL-1999; 99WO-US015810.
XX 14-JUL-1998; 98US-0092833P.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX Allen SM, Falco SC, Thorpe CJ;
XX WPI; 2000-195025/17.
XX P-PSDB; AAY44936.
XX Nucleic acid fragments encoding sulfate assimilation proteins in plants
XX and seeds useful as probes for isolating cDNAs and genes encoding
XX homologous proteins, in producing transgenic plants.
XX Claim 3; Page 41; 79pp; English.
XX The present sequence is a cDNA clone encoding corn sulphate permease, a
XX sulphate assimilation protein. This sequence is from a contig composed of
XX clones crin.pk0015.a2, p0006.cbys25rb, p0072.commc25r, p0091.cmard29r
XX and p0092.chwa43r from crin, p0006, p0072, p0091, and p0092 libraries,
XX respectively. This sequence is used as a probe to isolate other plant
XX sulphate assimilation proteins, for genetic and physical mapping of
XX related genes and as markers of traits linked to the gene. This is useful
XX for plant breeding and to construct chimeric genes, used to create
XX transgenic plants with altered levels of sulphate permease. The sulphate
XX permease peptides are useful for producing antibodies, that are used to
XX screen and isolate cDNA clones
XX
SQ Sequence 1981 BP; 388 A; 621 C; 612 G; 360 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,86e-213 Length: 1981
Score: 2164.00 Matches: 403
Percent Similarity: 86.09% Conservative: 86
Best Local Similarity: 70.95% Mismatches: 79
Query Match: 62.27% Indels: 0
DB: 3 Gaps: 0

US-10-762-049-18 (1-680) x AA250482 (1-1981)
QY 105 LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer 124
DB 8 GAGTCGACCTGATCGCGGCATCACCATCGCCAGCTCGCCATCCGACGGGCATCAGC 67
QY 125 TyrAlaLysLeuAlaIleLeuProProIleLeuGlyLeuTyrSerSerPheIleProPro 144
DB 68 TAGCCCAAGCTCGCCAAACCTGCGCGCGGTGCTCGGACTCTACTCGAGCTTGGTGGCGCG 127
QY 145 LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly 164
DB 128 CTGGTGTACCGCGTGTATGGGGAGCTCCAAGAGCTCGCGGTGGGACGGTGGCGGTGGCG 187
QY 165 SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlnGlyProLysLeu 184
DB 188 TCCTGCTCATAGCTCCATGCTCGCAGCGAGGTGTGCGCCGACGAGAACCCCGTGTCTC 247
QY 185 TyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuGly 204
DB 248 TACCTGCACCTCGCCCTTACCGGCACCTTCTTCGCGGGGTCTTCCAGGCTCGCTCGGC 307
QY 205 LeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMet 224
DB 308 CTCTCAGGTGGGCTTCACTCGTGGACCTGCTCGCACGCGACGATCGTGGGGTTCATG 367
QY 225 GlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis 244
DB 368 GCCGGCGCGCGACGGTGTGTGCTTCAGCAGCTGACGAGGATGTCGGGCTCTGTCCAC 427
QY 245 PheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlu 264
```

```
DB 428 TTCACCACTCCACCGAGCGTCTCCGTATGGAATCGTCTTACGACGACACACACAG 487
QY 265 TrpArgTrrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuSerThrArg 284
DB 488 TGGCGGTGGGAGAGCGTCTCGTTCGGCTCGCGCTTCTCTTCTTCTCTCTCTCTCT 547
QY 285 TyrPheSerLysLysArgProArgPhePheTrrpValSerAlaMetAlaProLeuThrSer 304
DB 548 TTCATCAGCAAGAGCGTCCCAAGCTGTTCGTGATCTCCGCGCGCGCGCGTGTGAGCTC 607
QY 305 ValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyValGluVal 324
DB 608 GTCTGTCTCGGAGCGTCTCGTGTACCTCACGCGCTGAAACACACGCGCATCGAAGTG 667
QY 325 IleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSer 344
DB 668 ATCGGTACTCTGAAGAAAGCGCTGAATCCACGCTCGGTCAAGCGCTGCAATTCCTCACCG 727
QY 345 ProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGlu 364
DB 728 CCTATCATGATCTCTCGGCTCAAGACTGGGATCATCACCGGCTCATTTGCCCTCGCCGA 787
QY 365 GlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLys 384
DB 788 CGAATCGCGCTGGGAGGAGCTTCGCCATGTTCAAGAACTACCATGATCGGACGACAAAG 847
QY 385 GluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThr 404
DB 848 GAGATGATCGCATCGGACGATGAACGTCCTGGGCTCGCTCACGTCGTCGTACCTGACG 907
QY 405 ThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSer 424
DB 908 AGGGGGCTTCTCGGCTCCGCGCTGAACTACACGCGCGGTGACGAGCGCCATGTCTG 967
QY 425 AsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHis 444
DB 968 AACGTGGTTCATGCTCGTGGCGGTGATGGTTCAGCTGCTTCTGACGCGCTGTTCAC 1027
QY 445 TyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAsp 464
DB 1028 TACACGCGCTCGTGTGCTGCGCGATCATGCTCTCCGCGATGCTGGGCTGCGCTGCG 1087
QY 465 TyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSer 484
DB 1088 TTCGGGCGCGCTGCACCTGTGGCGCTCGACAGGTGCGACTTCTGCTCTGCGCGCGC 1147
QY 485 AlaTyrIleGlyValPheGlySerValGluIleGlyLeuValIleAlaIleValIle 504
DB 1148 GGTACTCTGGCGCTCGTCTTCGCGCAGCGTTCGAGGTGCGGCTGCTGCTGCTGCGCTC 1207
QY 505 SerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIle 524
DB 1208 TCCTGTCTCGCGTCTCTGCTGTGCTGCGCGCGCGCGACGACGACGCTGCTGCGCAATC 1267
QY 525 ProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGly 544
DB 1268 CCGGACCATCGTGTACCGGAGGTGGACAGTACGCGCGCGCGCAGAGCGTGTGCGCGC 1327
QY 545 MetLeuLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGlu 564
DB 1328 GTGCTGTGCTGCGGCTCGACGCGCGCTCTACTTCGCCAAGCGCAGCTTACCTGGAGAG 1387
QY 565 ArgIleThrArgTrrpIleAspGluGluArgGlyIleLysAlaThrGlnThrSer 584
DB 1388 AGGATCTCGCGTGTATCGACGAGGAGGAGCGCCACAGAGCGAGCGCGGAGATGGGC 1447
QY 585 LeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSer 604
DB 1448 GTGCGGTAGTGTCTCTCGACATGGGTGCCATCGGTAGCATCGACACGAGCGGAGCAGC 1507
QY 605 MetLeuGluValLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsn 624
```

Db	1508	ATGCTGGACGAGCTCAACAGTCTTGGACAGAGGGGAATGCAGATCGTCTGGCGAAC	1567	PR	04-JUN-1999;	99US-0137502P.
				PR	07-JUN-1999;	99US-0137724P.
Qy	625	ProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLys	644	PR	08-JUN-1999;	99US-0138094P.
		::: :::				

PR	06-AUG-1999;	99US-0147303P.	Score:	1902.50	Matches:	354
PR	06-AUG-1999;	99US-0147416P.	Percent Similarity:	76.8%	Conservative:	125
PR	09-AUG-1999;	99US-0147493P.	Best Local Similarity:	56.82%	Mismatches:	143
PR	09-AUG-1999;	99US-0147935P.	Query Match:	54.75%	Indels:	1
PR	10-AUG-1999;	99US-0148171P.	DB:	3	Gaps:	1
PR	11-AUG-1999;	99US-0148319P.	US-10-762-049-18 (1-680) x AAC42229 (1-2006)			
PR	12-AUG-1999;	99US-0148341P.	QY	36	AsnAsnPheGluArgValHisGlnValProProGlnProPhePheLysSer	55
PR	13-AUG-1999;	99US-0148565P.	DB	102	AACCAATTAATGGAGTGCACAGGTGGTTCGCCGCATAGAGCAGGTGGCAAG	161
PR	16-AUG-1999;	99US-0148684P.	QY	56	LeuLysTyrSerLeuLysGluThrPhePheProAspAspProLeuArgGlnPheLysAsn	75
PR	17-AUG-1999;	99US-0149175P.	DB	162	CTAAAGACAAAACCTGAAGGAGACTTTCTTCCCGGACGATCCCTTAAGACAGTTCCAGAGCA	221
PR	18-AUG-1999;	99US-0149426P.	QY	76	LysProAlaSerLysLysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrp	95
PR	20-AUG-1999;	99US-0149722P.	DB	222	CAACCAAAACCGTACCAGCTCATACGCCGCTCAATATACATTTTCCCAATCTCCAATGG	281
PR	20-AUG-1999;	99US-0149723P.	QY	96	AlaProLysTyrThrPheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAla	115
PR	22-AUG-1999;	99US-0149929P.	DB	282	TGTCCTGAGTACAGCTTCAGCCTACTCAATCCGAGCTGTTTCAGGCTTTACCATCGCT	341
PR	23-AUG-1999;	99US-0149920P.	QY	116	SerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeu	135
PR	23-AUG-1999;	99US-0149930P.	DB	342	AGTTTAGCTATTCTCTCAGGGATTAGTTACGGGAAGCTAGCAAAATTTGCCCAATCGTT	401
PR	25-AUG-1999;	99US-0150366P.	QY	136	GlyLeuTyrSerSerPheIleProLeuIleTyrAlaMetMetGlySerSerArgAsp	155
PR	26-AUG-1999;	99US-0150884P.	DB	402	GGTCTATATCTCAGCTTCGTCACCGTGGTGGTTATGCGGTCTTGGGAAGCTCAAGAGAT	461
PR	27-AUG-1999;	99US-0151065P.	QY	156	LeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAla	175
PR	27-AUG-1999;	99US-0151066P.	DB	462	CTAGCAGTGGGACCAAGTCTCCATAGCGTGGTAAATCTTAGGATCCATGCTAGGCAACAA	521
PR	30-AUG-1999;	99US-0151303P.	QY	176	ValAspProAsnGluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPhe	195
PR	31-AUG-1999;	99US-0151438P.	DB	522	GTATCTCCGTTAGACGATCCTCTTCTTACAGCTAGCCTTCTCTTCTTCTTCTTCTT	581
PR	01-SEP-1999;	99US-0151930P.	QY	196	AlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeu	215
PR	07-SEP-1999;	99US-0152363P.	DB	582	GCTGTCTCTTTCAAGCCTCTCTTGGAAATCTCAGCGTGGGATTTATATAGACTTTCTA	641
PR	13-SEP-1999;	99US-0153758P.	QY	216	SerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGln	235
PR	15-SEP-1999;	99US-0154018P.	DB	642	TCAAAAGCGACGCTAATAGGGTTATGGGTGAGGAGCCATAAATGTATCATCTCCAACAG	701
PR	16-SEP-1999;	99US-0154039P.	QY	236	LeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMet	255
PR	20-SEP-1999;	99US-0154179P.	DB	702	CTAAAGGGTCTGCTTGGGATAACTCATTTCAAAAAGCATATGAGTGTAGTCTGTTCTC	761
PR	22-SEP-1999;	99US-0155139P.	QY	256	ArgSerValPheThrGlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysVal	275
PR	23-SEP-1999;	99US-0155486P.	DB	762	TCCTCTGTTTTCCAAACACACCAACGAGTGGTCATGGCAACCAATGTGATGGAGTTGCG	821
PR	24-SEP-1999;	99US-0156458P.	QY	276	PheIlePhePheLeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrp	295
PR	29-SEP-1999;	99US-0156596P.	DB	822	TTCTTGCTCTTCTTGCTCTCGACAGTCACCTCAGCATGAAGAAGCCGAGCTGTTTGG	881
PR	04-OCT-1999;	99US-0157117P.	QY	296	ValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThr	315
PR	05-OCT-1999;	99US-0157753P.	DB	882	GTCTCAGCGGAGCACCACTTCTTCCGTTATCGTCTCTACACTTCTTGTCTTGTCTTC	941
PR	06-OCT-1999;	99US-0157865P.	QY	316	HisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysLysLeuAsnProPro	335
PR	07-OCT-1999;	99US-0158029P.	DB	942	AGAGCCGAGCGTCACGGAATCAGCGTCATCGGGAATTTACGAGAAGGTTTGAATCCACCG	1001
PR	08-OCT-1999;	99US-0158232P.	QY	336	SerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIle	355
PR	12-OCT-1999;	99US-0158369P.	DB	1002	TCTTGGAAACATGCTTCAGTTTCACCGTAGTCATCTCGCAGCTCGTCCCAAAACCGACTC	1061
PR	13-OCT-1999;	99US-0159293P.	QY	356	ValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyr	375
PR	13-OCT-1999;	99US-0159294P.				
PR	13-OCT-1999;	99US-0159295P.				
PR	14-OCT-1999;	99US-0159329P.				
PR	14-OCT-1999;	99US-0159330P.				
PR	14-OCT-1999;	99US-0159331P.				
PR	14-OCT-1999;	99US-0159637P.				
PR	14-OCT-1999;	99US-0159638P.				
PR	18-OCT-1999;	99US-0159584P.				
PR	21-OCT-1999;	99US-0160741P.				
PR	21-OCT-1999;	99US-0160767P.				
PR	21-OCT-1999;	99US-0160768P.				
PR	21-OCT-1999;	99US-0160770P.				
PR	21-OCT-1999;	99US-0160814P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.				
PR	22-OCT-1999;	99US-0160981P.				
PR	22-OCT-1999;	99US-0160989P.				
PR	25-OCT-1999;	99US-0161404P.				
PR	25-OCT-1999;	99US-0161405P.				
PR	25-OCT-1999;	99US-0161406P.				
PR	26-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	28-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161922P.				
PR	28-OCT-1999;	99US-0161993P.				
PR	29-OCT-1999;	99US-0162142P.				

Alignment Scores:
Pred. No.:

4.36e-186

Length:

2006

Db	248	CTTGCCAACTGGCCACCATTCTTTGGGATTAATTCGAGTTTGTTCGCCCAATGATATAC	307
Qy	148	AlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeu	167
Db	308	TCGCTGCTTGAAGTTCTAGACATCTTGGTGTGGACCTGTTTCCATTGGCTTTGGTC	367
Qy	168	MetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHis	187
Db	368	ATGGGATCAATGTTAAAGTGATAAAATTTCTTACACTCAAGAACCTATTTCTATCTGGGA	427
Qy	188	LeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArg	207
Db	428	TTGCTTTTACCCGCCACTTCTTTGCTGGTGATTTCCAAGCTCTCTGGGTATATAAAGG	487
Qy	208	LeuGlyLeuLeuLeuAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAla	227
Db	488	CTAGGCTTCGTAATGATTTTCTGTGAGGCAACGCTGGTGTGATTCACAGGGCGTGCT	547
Qy	228	AlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHis	247
Db	548	GCCATTATTGTGTCACCTGCAGCAGCTGAAAGTTTACTTTGGAATAGTGCACTTTACCAGC	607
Qy	248	GlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGluTyrArgTyr	267
Db	608	AAGATGCAAAATAATTCAGTAACGATCTCTGTTTCAAGCAAAACACAGCAGTGTCTATGG	667
Qy	268	GluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSer	287
Db	668	CAAAACCATTTTGGGATTCGGCTTCTGGTCTCTTGTCTGTCGACAAACAGCAGCATTAGT	727
Qy	288	LysLysArgProArgPhePheTyrValSerAlaMetAlaProLeuThrSerValIleLeu	307
Db	728	TTGAGGAACCAAAACTATTCTGGTTCACGACCTGCCCATTTGACATCAGTTATTCTG	787
Qy	308	GlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGlu	327
Db	788	TCAACCATTTTAGTCTTTCTCTGAGAAATAAGACTCATCAAAATTTTCAGTTATTGGGCAC	847
Qy	328	LeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMet	347
Db	848	TTACCAAGGGAGTTAATCCACCATCAGCAAAACATGTTATACTTCAATGGTCTTACTGT	907
Qy	348	ThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAla	367
Db	908	GGTCTGCTATCAAACTGGCATCATCACAGGATCTTATCTCTCCTCAGGAAATGCA	967
Qy	368	ValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIle	387
Db	968	GTAGGAGAACATTTGCTTCTACTTTAAGAACTTACCAGGTGGATGGAAACAAAGAAATGATG	1027
Qy	388	AlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyPro	407
Db	1028	GCCATTGGTCTAAATGAACATGATGGCTGGTGTCTTTCATGTTATGTTTCAACGGGATCC	1087
Qy	408	PheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIle	427
Db	1088	TTTTCTCGATCGGTGTTAATAATGCTGGACACAGACACAGCTTCAATATATATC	1147
Qy	428	MetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrPro	447
Db	1148	ATGGCTGCAGCTGTCTAGTGACACTTCTGTTTCTCATGCTCTTCTACTATACACCA	1207
Qy	448	LeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAla	467
Db	1208	AATGTTGCTTTAGCGGCCATTTATCATCAGTGTGTGATGGTCTTAATAGATATATCAATC	1267
Qy	468	AlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIle	487
Db	1268	GCATATAAATTGTGGAAGGTTGACAAACTGATTTCTTGGCTGTGTGTGCTCCTTTTTT	1327
Qy	488	GlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValIleu	507

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 6; SEQ ID NO 528; 899pp; English.
PS
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC

CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (MI) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX Sequence 1896 BP; 483 A; 465 C; 426 G; 522 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.04e-185 Length: 1896
 Score: 1898.50 Matches: 353
 Percent Similarity: 77.35% Conservative: 125
 Best Local Similarity: 57.12% Mismatches: 139
 Query Match: 54.63% Indels: 1
 DB: 8 Gaps: 1

US-10-762-049-18 (1-680) x ADA68554 (1-1896)

Qy 41 ValHisGlnValGluValProProGlnProPhePhePheLysSerLeuLysTyrSerLeu 60
 Db 7 GTGCACAAAGGTGGTGTCTCCGCGCATAGAGCAGCGGTGCAAAAGCTAAAGACAAAACTG 66
 Qy 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
 Db 67 AAGGAGACTTCTTCCCGGACGATCCCTTAAGACAGTTACAGAGCAACCAACCGGTACC 126
 Qy 81 LysPheMetLeuGlyLeuGlnPhePhePheProLlePheGluTrpAlaProLysTyrThr 100
 Db 127 AAGCTCATACGCGCGCTCAATACATTTTCCATCTCCATGGTGTCTCTAGTACAGC 186
 Qy 101 PheGlnPheLysAlaAspLeuAlaGlyLeuAlaGlyLeuAlaSerLeuAlaLlePro 120
 Db 187 TTCAGCCTACTCAAAATCCGACGCTGTTTACAGTCTTACCATCGTAGTTTGTCTTCTT 246
 Qy 121 GlnGlyLeuSerTyrAlaLysLeuAlaAsnLeuProProLleLeuGlyLeuTyrSerSer 140
 Db 247 CAGGGGATAGTAGTCGCGAAGCTAGCAAAATTTGCCAACATCTGTTGTTCTATCTGAGC 306
 Qy 141 PheileProLeuLeuTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
 Db 307 TTCGTGCCACCGTGGTTATGGTCTGGGAGCTCAAGAGATCTAGCAGTGGGACCA 366
 Qy 161 ValAlaValGlySerLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
 Db 367 GTCTCCATACGTCGTTAATCTTAGGATCCATGTAAGGCCAACAGTATCTCCCGTAGAC 426
 Qy 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
 Db 427 GATCCTGTCTCTTCTACAGCTAGCCCTCTCTTCACTTCTTGTGCTCTTTTCAA 486
 Qy 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuAlaValAspPheLeuSerHisAlaThrIle 220
 Db 487 GCCTCTCTGGAATCTCAGCTGGGATTTATATAGATCTTCTATCAAAAGCAGCGCTA 546
 Qy 221 IleGlyPheMetGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
 Db 547 ATAGGGTTTATGGTGAGCAGCCATAATGTATCTCACTCCACAGCTAAGGGTCTGCT 606
 Qy 241 GlyLeuGlnHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThr 260
 Db 607 GGGATAACTCATTTTCAAAAGCATATAGTGTAGTCCCTCTCTCTCTCTCTCTCTTCAA 666
 Qy 261 GlnThrHisGlnTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280
 Db 667 CACACCAAGGAGTGTATGGCAACATTTGATGGAGTTTGTCTCTCTCTCTCTCTCTCT 726
 Qy 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
 Db 727 CTCTCGACACGTCACTCAGCATGAGAACCGGAGCTGTTTGGGTCTCAGCGGAGCA 786
 Qy 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320

Db 787 CCACCTTCTTCCGTTATCGTCTCTACACTTCTTGTCTTTTTCAGAGCGGCGTCAC 846
 Qy 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProSerLeuThrAsnLeu 340
 Db 847 GGAATCAGCTCATCGGAAATTTACAGAAAGGTTGAATCCACCGTCTTGAACATGCTT 906
 Qy 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyLeuValValGlyIle 360
 Db 907 CAGTTTCACGGTAGTCTATCGCACTCGTCCCAAAACCGGACTCTGTCACGGAATCGTC 966
 Qy 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
 Db 967 TCCCTCAGCGAAGAAATCGAGTGGGAAGAAATTTGACGCGTAAAGAAATCACCAGTA 1026
 Qy 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
 Db 1027 GATGGAACAAAGAGATGATCGCATTTGCTGATGACGATAGTGGCTCTGCCACATCT 1086
 Qy 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
 Db 1087 TGCTAGCTCACAAACCGGAGCATTTCTTAGATCAGCGGTCAACAAACACGCGGAGCTAAA 1146
 Qy 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
 Db 1147 ACCGCAAGTTTCAACATTTGTTATGTCGGTCACTGTTATGTTACGCTTCTTCTCTAATG 1206
 Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
 Db 1207 CCGCTTTTGAATACACTCCCAATGGTCTCTCGTGCCATCATTTGTACCGGGCTCAT 1266
 Qy 461 GlyLeuLeuAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
 Db 1267 GGTCTCATCGACCTTCCCGCGGCTGTCAATATGGAAGATCGATAAATTTGATTTCTTG 1326
 Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
 Db 1327 GTGATCTTTCGCGGTCTTGTGTCATTTTCTTCTTCCGTTCAAAACCGTCTAGCCATT 1386
 Qy 501 AlaIleValIleSerValLeuValLeuPheIleAlaArgProArgThrPheVal 520
 Db 1387 GCGGTGGGCTATCGTTGTTCAAGATTTGATGCAAGTAACAAGCGGAAATGGTTATA 1446
 Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
 Db 1447 ATGGGTAAATCTCTGGAACGGATATATACCGAGATCTTCTCATTTACAAAGAACACAA 1506
 Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
 Db 1507 AAGATCCCGGATTTCTGTTTAAAGCATCGAATCTCTGTCATTTTCCCAATTTCTAAC 1566
 Qy 561 TyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThr 580
 Db 1567 TACCTCAGTGAAGAACATCTCTGTTGATTAAGAAATGCGAAGAGGAGGAGTCAAGAG 1626
 Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
 Db 1627 AAGCAATCTAGCTACAGTCTTCTGATTTCTGAAATGTCAGCGGTGAGGAGTACACAC 1686
 Qy 601 SerGlyIleSerMetLeuGluGluValLysIleThrGluArgGluLeuGlnLeu 620
 Db 1687 AACGGAGTGTCTTTTTTAAGGAACCTTAAGAAACCAACCCGCAAGGAGGAGCATCGAGCTT 1746
 Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
 Db 1747 GTGTTTGTGAACCTCTTAAGCGAGTGGGAGAGCTTCAAGAGCTGACGACAAACAAA 1806
 Qy 641 His---LeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAla 657
 Db 1807 GAGTTTCATGAGCCCGAGTCTTCTTCTTAAACCGTCTGCTGAGGCGCTGCGTCG 1860

RESULT 11
 AA250481

ID AAZ50481 standard; cDNA; 2279 BP.
 AC AAZ50481;
 DT 23-MAY-2000 (first entry)
 XX Corn sulphate permease-1 cDNA clone.
 KW Sulphate Permease; sulphate assimilation protein; corn; contig; probe;
 KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
 KW antibody; screen; ss.
 XX
 OS Zea mays.
 XX
 XX Key Location/Qualifiers
 XX CDS 3..2069
 FT /*tag= a
 FT /product= "Corn sulphate permease-1"
 FT /note= "Derived from a contig composed of clones
 FT cbn10.pk0062.b10, ccoln.pk081.h21, ccoln.pk092.12,
 FT csc1c.pk005.j3, p0004.cblej58r, p0089.csdch19r,
 FT p0094.csssg12r, p0121.cfrmx30r and p0128.epicz09r"
 XX
 PN WO200004154-A2.
 XX
 XX 27-JAN-2000.
 XX
 XX 13-JUL-1999; 99WO-US015810.
 XX
 XX 14-JUL-1998; 98US-0092833P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Allen SM, Falco SC, Thorpe CJ;
 XX
 XX WPI; 2000-195025/17.
 DR P-PSDB; AAY44935.
 XX
 XX Nucleic acid fragments encoding sulfate assimilation proteins in plants
 PT and seeds useful as probes for isolating cDNAs and genes encoding
 PT homologous proteins, in producing transgenic plants.
 XX
 PS Claim 3; Page 38-39; 79pp; English.
 XX
 CC The present sequence is a cDNA clone encoding corn sulphate permease, a
 CC sulphate assimilation protein. This sequence is from a contig composed of
 CC clones cbn10.pk0062.b10, ccoln.pk081.h21, ccoln.pk092.12, csc1c.pk005.j3,
 CC p0004.cblej58r, p0089.csdch19r, p0094.csssg12r, p0121.cfrmx30r and
 CC p0128.epicz09r from cbn10, ccoln, csc1c, p0004, p0089, p0094, p0121 and
 CC p0128 libraries, respectively. This sequence is used as a probe to
 CC isolate other plant sulphate assimilation proteins, for genetic and
 CC physical mapping of related genes and as markers of traits linked to the
 CC gene. This is useful for plant breeding and to construct chimeric genes,
 CC used to create transgenic plants with altered levels of sulphate
 CC permease. The sulphate permease peptides are useful for producing
 CC antibodies, that are used to screen and isolate cDNA clones
 XX
 SQ Sequence 2279 BP; 512 A; 606 C; 629 G; 532 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 5,26e-180 Length: 2279
 Score: 1844.50 Matches: 355
 Percent Similarity: 74.25% Conservative: 115
 Best Local Similarity: 56.08% Mismatches: 158
 Query Match: 53.08% Indels: 5
 DB: 3 Gaps: 3
 XX
 US-10-762-049-18 (1-680) x AAZ50481 (1-2279)
 QY 41 ValHisGlnValGluValProProGlnProPhePheLeuLysSerLeuLysTyrSerLeu 60
 DB 189 GTGCACAAGTGGCGCGCCACCGCGCGGAGCACCGGACGACGAGATGAAGGTGAGGGTG 248

QY 61 LysGluThrPhePheProAspPheProLeuArgGlnPheLysAsnLysProAlaSerLys 80
 DB 249 AAGGAGACCTTCTCCCGCAGCAGCCGTTCCGGCGGTTCAGGGGCGAGCGCGGGGACG 308
 QY 81 LysPheMetLeuGluLeuGlnPhePhePheProIlePheGluTyrPheProLysTyrThr 100
 DB 309 CAGTGGCTCATGGCGGTACGTACCTCTTCCCATCTCTGGACTGGGTGGCGAGCTACTCC 368
 QY 101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
 DB 369 TTGTGCTCTTCAAGTCCGACCTCGTGGGGGCTCCACCATTCGCCAGCTCCGCTCTCT 428
 QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSer 140
 DB 429 CAGGGCATTTAGCTACGCGAAGCTGGCAAGCTTGCCTCCCATTAATCGGGCTGTATTCGAGC 488
 QY 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
 DB 489 TTCTGTCGCCCATGGTGTACGCGGTGTGGGAGCTCCCGTACCTGGCGGTGGGCGCG 548
 QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
 DB 549 GTGTGATCTCGTGTATCATGGGTCCTCATGCTGGCGAGCGCTGAGCCCCACTGGCG 608
 QY 181 AspProLysLeuTyrIleHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
 DB 609 GAGCCGACGCTGTCTCTGACGCTGGCTTCACTCCACCCTGTTCGCGGGGTGGTGACG 668
 QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
 DB 669 GCCTCCTCGGCTATCTCAGGCTCGGCTTCGTCATCGACTTCTGTCCAAAGCGAGCTG 728
 QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
 DB 729 GTGGGTTTCATGCGCGCGCGCCCATCATCTGCGGGCTGCACAACTCAAGGGGCTGCTG 788
 QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
 DB 789 GGCATCGTCCACTTCAACACCGAGATGGCATCGTCCAGTCATGGCTCCGCTTCCAC 848
 QY 261 GlnThrHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
 DB 849 CACACCGCGAGTGTGTGGCGAGCATCTCTATGGGCGTCTGCTTCTCTCTCTCTCTG 908
 QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTyrPheValSerAlaMetAla 300
 DB 909 CTGTGGCGAGCATGTGAGCATCATAGTGGCCAAAGCTTTCTGGGTTTCGGCGTGGCGG 968
 QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGlyLysHis 320
 DB 969 CCCCCTGGCATCGTCAACCATCTCGACGCTGCTTTTCTCTCTTCAAAAGCTCAGAACCAT 1028
 QY 321 GlyValGluValIleGlyLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
 DB 1029 GGCATCAGCATCATTTGGCGAGCTCAAGTGGCGCTGAAATCGCCCTCGTGGGACAGCTC 1088
 QY 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
 DB 1089 CFTTTTGACACGGCGTATTAGCGCTCCACATGAGACTGGCGCTTGTCCCGGATCATC 1148
 QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
 DB 1149 TCACCTGACGGAAGGAATAGCGGTGGTAGAACATTTGCTCTCACTCAAGGACTACCAGATA 1208
 QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
 DB 1209 GATGGAACAAGGAGATGATGGCCATAGGTTGATGAATGTTGTTGGGTCTCGCATCA 1268
 QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
 DB 1269 TGCTAGCTAACCAACAGGTGCGTTCTCCCGCTCTGTGTAAACACACACGCCGCGCTCAAG 1328
 QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuLeuThrLeuPheLeuThr 440


```
Db 1329 ACTGCCATGTCACACGGTATCATGGCGTCACTGTGATGGTCACGCTGCTTCCTCATG 1388
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleValSerAlaMetLeu 460
Db 1389 CCACTGTTCTGTATACACACCAAGTTGTCTCGAGCGATCATCATCGCGGGTATC 1448
Qy 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 1449 GGCGTGTATCATTTCCCGGGGTGTACCATCTGGAAGATGACACAGATGGATTTCTG 1508
Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db 1509 GTGTGCGTTTGGCGGTTCGGCGGTCTTCTCATCTCACTCAAGAACGCTTGCATTA 1568
Qy 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
Db 1569 GCGGTTGGTATATCTATATTAGGTGTGTGATGCATCACAGGCCGAAGATGATGGTT 1628
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1629 CAAGGGAACATCAAGGGAGCTGATATTATACAGAGACCTGCTCACTACAAAGGAGGCCAA 1688
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db 1689 AGAGTTCTGGTCTTGATCTTGGCCATTGAGCACCCGATAAATTCGCCAACTCCAAC 1748
Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db 1749 TACTGTAATGAAGATTAAAGATGGATA-----GAGGAAGAATCTTTTGAACAGGAT 1802
Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db 1803 AAACATACCTGAATCCCATTTTCATAATCTGGATCTGTCACTGTCTCTGCAATTGACACA 1862
Qy 601 SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db 1863 AGTGGCATAGCGTTCTCATGTGACATAAAGAAATCAATAGAGAAACGTGTCTGGAGCTT 1922
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLys---SerLysPheGln 639
Db 1923 GTGCTGTGCAATCAACTGGAGAGTCAATGAGAAATAACAAGTGCACAAAGGCTGAA 1982
Qy 640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db 1983 AACTATTTTAGGCGCAGATTGCTTGATCTGACCACTGGCGAAGCAATC-----GCTTCA 2036
Qy 660 PheAsnLeuArgAlaSerLysThrAsnProLysLysAsp 672
Db 2037 CTTTCTGCACCTTGCCAAGATGCACAAACCCCTAAATGGAT 2075

RESULT 12
ADA71315
ID ADA71315 standard; DNA; 1940 BP.
XX
AC ADA71315;
DT
XX 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 4638.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
XX W0203000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
```

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

XX Claim 6; SEQ ID NO 4638; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

XX Sequence 1940 BP; 535 A; 421 C; 432 G; 552 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,89e-177	Length:	1940
Score:	1817.00	Matches:	341
Percent Similarity:	72.21%	Conservative:	119
Best Local Similarity:	53.53%	Mismatches:	163
Query Match:	52.29%	Indels:	14
DB:	8	Gaps:	3

US-10-762-049-18 (1-680) x ADA71315 (1-1940)

Qy 34 GlyMetAsnAsnPheGluArgValHisGlnValProProGlnProPhePhe 53
Db 67 GGATCATCTCGAAGAACCGGTGTGTACAGGTGAGGTCTCTCTCAAAAGAACTTTGCG 126

Qy 54 LysSerLeuLysTyrSerLeuLysGluThrPhePheProAspAspProLeuArgGlnPhe 73
Db 127 AGAGAGTTTAGAGACACATTTGAGAGAGACTTCTTCCATGACAAACCATTTACCCAGTAC 186

Qy 74 LysAsnLysProAlaSerLysLysPheMetLeuGlyLysGlnPhePheProIlePhe 93
Db 187 AAGGACCAATCTGGATCGGCAAGTTCAATGATGGCGCTGCAGTCTCTGTCTCCAACTTTT 246

Qy 94 GluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThr 113
Db 247 GAGTGGGCGAGGTGCTACAACTTAAGAAAGTTCAAGAGTGATTTAATTGCTGATTGACC 306

Qy 114 IleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuPro 133
Db 307 ATTGAAGTCTCTGCATTTCCCGAGGATATCGGGTATTTCGAAGCTTCTTAACCTTGGATGCA 366

Qy 134 IleLeuGlyLeuTyrSerPheIleProLeuIleTyrAlaMetMetGlySerSer 153
Db 367 CAATATGGACTATACCTCCAGCTTTGTCCCAATGATATATGCTCAATGGCAGCTCA 426

Qy 154 ArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMetLeuSer 173
Db 427 AAGGATATTGCAATCGTCCGGTTGCAGTG----- 456

Qy 174 AsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThr 193
Db 457 AATGAAGTCGACCTGTCTCAAAAACAGAGGAATATCTGCGCTTCTGCTTTCACGGCTACC 516

Qy 194 LeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuLeuValAsp 213
Db 517 TTCCTTTGGGTATCACTCAAGCAGCCCTAGGATTTTAAAGGCTAGGATTCCTCTCATCGAG 576

Qy	214	PheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValValCysLeu	233
Db	577	TTCTTGTCACATGCTGCAAAATGTTGGGTTTATGGAGGAGCAGCCATTACTATTGGCCCTT	636
Qy	234	GlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAspIleIleSer	253
Db	637	CAGCAGCTAAATAATATGTGTGGGCATCAAGAGCTTTTACAAAGAAAACCGATATAATTTCT	696
Qy	254	ValMetArgSerValPheThrGlnThrHisGlu--TrpArgTrpGluSerAlaValLeu	272
Db	697	GTATTGCGCTCTGTTTGGACTTCAGCTCATCCGGGTGGAATTCGCAAACTATTGTGATT	756
Qy	273	GlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSerLysLysArgProArg	292
Db	757	GGCATAAATCTTCCVTGCAATTCCTTCTGCTGCCAAGTACATTGGAAAGAGAATAAGGAAG	816
Qy	293	PhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuVal	312
Db	817	TTCTTTCGGGTGCCAGCTATTGCTCTATATACTTCAGTTATTTTGGCAACCCCTTTTGTG	876
Qy	313	TyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeu	332
Db	877	TTCAATTACTCGTGTGACAGCAAGAGTGTTCAGATTGTTAACCCACATCAAAAAGGGCATA	936
Qy	333	AsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLys	352
Db	937	AAACCATCATCAGTCCCAAAATTTATTTTCACTGGTGCATTGTTGTGCAAAAGGTTTCAAG	996
Qy	353	ThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPhe	372
Db	997	ATCGGTGTCTATTCGGCCATGATCGGTTTAAACGGAAGCTGTGGCAATTTGGGAGGACGTTT	1056
Qy	373	AlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMet	392
Db	1057	GCTGCTCTGAAGGACTATCAATTAGATGGAAAACAAGGAGATGGTAGCATTGGAACATATG	1116
Qy	393	AsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAla	412
Db	1117	AAACATACAGGGTCAATGACATCTTGTATTATTCACACAGGTTCTTCTCTCGTTCGCGCA	1176
Qy	413	ValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaVal	432
Db	1177	GTCAATTTTCATGGCTGGCTGCCAAACACCAGTGTCCATATCATATTATGTGACGAGTTGTG	1236
Qy	433	MetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSer	452
Db	1237	CTCTGACCTTGTGGTTCATCACACCACTCTTCAAGTACACACCAAAATGGATCCTCGGG	1296
Qy	453	AlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPhe	472
Db	1297	TCGATAATCATTTCTCGGGTGATCGGGCTTGTGGACTACGAAGCAGTAATTTCTCATTTGG	1356
Qy	473	LysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPheGly	492
Db	1357	AAAGTTGCAAAAATGGATTTTCATTTCTTGATGGGAGCATTCCTTGGTGGGCTTCGCCA	1416
Qy	493	SerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeuPhe	512
Db	1417	TCTGTTGAGATAGGCTTTTAAATAGCTGTATCGATTTCATTTGCTAAAAATCTTCTTCAA	1476
Qy	513	IleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsn	532
Db	1477	GTAAACAGGCCAAGGACTGACTTCTCGGAAACCTTCAGGCACCAAAATATACAGGAAC	1536
Qy	533	ValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAla	552
Db	1537	ACTGATCAGTATCCAGAAGCAAGACATATTCCTGGGTTGTATAGTCAGGGTTGATTCT	1596
Qy	553	ProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGlu	572
Db	1597	GCTATCTACTTCTCCAACTCAAACTATGTCGGGAGAGAACTCTTATGGTGGTGCACGGAG	1656

Qy	573	GluGluGluArgIleIysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMet	592
Db	1657	GAAGAAGAGAAAGCCAAAGCAGAGGACAGCTCTAAAAATCAATTTCTTGATCATAGAGATG	1716
Qy	593	SerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValIysIle	612
Db	1717	TCCCGGTATAGATATCGATACAGTGGCATCCATGCTCTAGAGATCTATACAGAAT	1776
Qy	613	ThrGluArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetIysLys	632
Db	1777	CTTAAGAAGAGAGACATTCAGCTCATCTTAGCAAAATCCTGGATCCATCGTCATGATGAAG	1836
Qy	633	LeuAsnIysSerIysPheGlnAsnHisLeuGlyIysLysTrpIleTyrLeuThrValGlu	652
Db	1837	CTCTTGTCATCCCAAGCTCAACGAGCACAATAGGAGGATTAACAATATTTTCTCACGGTTGCT	1896
Qy	653	GluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerIysThrAsnPro	669
Db	1897	GAGCTGTGGTTCTGCACA-----AGAAAGTCGATGCAAGAACCA	1938
RESULT	13		
ABQ82719	ID	ABQ82719 standard; cDNA; 1971 BP.	
XX	AC	ABQ82719;	
XX	XX		
XX	DT	03-JAN-2003 (first entry)	
XX	XX	Arabidopsis thaliana sulphate transporter Sultr1;3 encoding cDNA SEQ.1.	
DE	DE	Arabidopsis thaliana; sulphate transporter; chromosome 1; plant;	
KW	KW	sulphur-containing metabolite; glutathione; phytoekratin; gene; ss.	
KW	XX	Arabidopsis thaliana.	
OS	OS		
XX	XX	Key Location/Qualifiers	
PH	PH	1..1971	
FT	FT	CDS	
FT	FT	/*tag= a	
FT	FT	/product= "sulphate transporter Sultr1;3"	
XX	XX	JP2002272472-A.	
XX	XX		
XX	PD	24-SEP-2002.	
XX	XX		
PF	PF	22-MAR-2001; 2001JP-00082891.	
XX	XX		
PR	PR	22-MAR-2001; 2001JP-00082891.	
XX	XX	(KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.	
PA	PA		
XX	XX	WPI; 2002-718707/78.	
XX	XX	P-PSDB; ABP53777.	
XX	XX		
XX	XX	A gene encoding sulfate ion transporter Sultr1;3, a polynucleotide, an	
PT	PT	oligonucleotide, a recombinant vector, a plant.	
XX	XX		
PS	PS	Claim 2; Page 6-9; 12pp; Japanese.	
XX	XX		
CC	CC	The present invention describes a gene which is present in the first	
CC	CC	chromosome of Arabidopsis thaliana and encodes sulphate ion transporter	
CC	CC	Sultr1;3 having the amino acid sequence given in ABP53777. Also	
CC	CC	described: a polynucleotide purified from the genomic DNA, the mRNA, the	
CC	CC	cDNA or their complementary sequences; an oligonucleotide hybridising	
CC	CC	with the above gene or the above polynucleotide and consisting of a base	
CC	CC	sequence of at least 10 base pairs; a recombinant vector carrying the	
CC	CC	above polynucleotide; and a plant in which the above gene is introduced	
CC	CC	or its descendant or their tissues. The plant can be used for the	
CC	CC	preparation of sulphur-containing metabolites such as glutathione and	
CC	CC	phytoekratin. The present sequence encodes Arabidopsis thaliana sulphur	
CC	CC	transporter Sultr1;3, from the present invention	
XX	XX		
SQ	SQ	Sequence 1971 BP; 504 A; 465 C; 427 G; 575 T; 0 U; 0 Other;	

Alignment Scores:																	
Pred. No.:	2.84e-176	Length:	1971														
Score:	1807.50	Matches:	335														
Percent Similarity:	72.76%	Conservative:	119														
Best Local Similarity:	53.69%	Mismatches:	169														
Query Match:	52.01%	Indels:	1														
DB:	6	Gaps:	1														
US-10-762-049-18 (1-680) x ABQ82719 (1-1971)																	
Qy	37	Asn	Phe	Glu	Arg	Val	His	Gln	Val	Pro	Pro	Gln	Pro	Phe	Phe	Leu	56
Db	76	AAT	AC	CGG	TAT	GC	CA	AA	GT	GC	CT	CT	CT	TA	AG	CA	135
Qy	57	Lys	Tyr	Ser	Leu	Gln	Val	Pro	Asp	Asp	Pro	Leu	Arg	Gln	Phe	Leu	76
Db	136	ATG	TAC	ACT	TTT	TA	AG	AA	CT	TT	CC	CA	GT	AT	CT	CT	195
Qy	77	Pro	Ala	Ser	Lys	Val	Phe	Met	Gly	Leu	Gln	Phe	Phe	Pro	Leu	Ala	96
Db	196	TCA	AA	TCC	AA	AA	AG	CT	AT	CG	GT	AT	CG	GT	AT	CG	255
Qy	97	Pro	Lys	Tyr	Thr	Phe	Gln	Phe	Leu	Leu	Ala	Asp	Leu	Ala	Gly	Leu	116
Db	256	AG	AA	AT	TA	AT	CT	TA	AG	TT	GC	GG	CA	CT	TA	AG	315
Qy	117	Leu	Ala	Leu	Pro	Gln	Gly	Ser	Tyr	Ala	Lys	Leu	Ala	Asn	Leu	Pro	136
Db	316	CT	CT	GC	AT	CT	CT	CA	GT	AT	GC	CA	GT	AT	GC	CA	375
Qy	137	Leu	Tyr	Ser	Ser	Phe	Leu	Pro	Leu	Leu	Tyr	Ala	Met	Met	Gly	Ser	156
Db	376	CT	AT	AT	TA	AT	CT	TA	AG	TT	GC	GG	CA	CT	TA	AG	435
Qy	157	Ala	Val	Gly	Thr	Val	Ala	Val	Gly	Ser	Leu	Met	Gly	Ser	Met	Leu	176
Db	436	GC	AG	TT	GC	AG	TT	GC	AG	TT	GC	AG	TT	GC	AG	TT	495
Qy	177	Asp	Pro	Asn	Glu	Asp	Pro	Lys	Leu	Tyr	Leu	His	Leu	Ala	Phe	Thr	196
Db	496	G	A	C	C	C	A	A	C	C	C	A	A	C	C	C	555
Qy	197	Gly	Val	Phe	Gln	Ala	Leu	Gly	Leu	Phe	Arg	Leu	Leu	Val	Asp	Phe	216
Db	556	GG	TG	TA	CT	CA	AG	CA	CT	CG	GA	TT	CT	CG	GA	TT	615
Qy	217	His	Ala	Thr	Leu	Leu	Gly	Phe	Met	Gly	Val	Ala	Ala	Thr	Val	Cys	236
Db	616	C	A	C	G	G	G	T	TA	CT	CA	AG	CA	CT	CA	AG	675
Qy	237	Lys	Ser	Leu	Gly	Leu	Gly	His	Phe	Thr	His	Gly	Ala	Asp	Leu	Ser	256
Db	676	AA	AG	GC	TT	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	735
Qy	257	Ser	Val	Phe	Thr	Gln	Thr	His	Glu	--	Trp	Arg	Trp	Glu	Ser	Ala	275
Db	736	T	C	G	T	A	T	A	T	A	T	A	T	A	T	A	795
Qy	276	Phe	Leu	Phe	Leu	Ser	Thr	Arg	Tyr	Phe	Ser	Lys	Lys	Arg	Pro	Arg	295
Db	796	T	T	C	T	G	A	T	T	C	C	A	G	T	T	T	855
Qy	296	Val	Ser	Ala	Met	Ala	Pro	Leu	Thr	Ser	Val	Leu	Gly	Ser	Leu	Val	315
Db	856	A	T	T	C	A	G	T	T	A	T	T	C	A	G	T	915
Qy	316	His	Ala	Gly	Lys	His	Gly	Val	Glu	Val	Leu	Gly	Lys	Lys	Gly	Leu	335
Db	916	C	G	A	C	C	G	A	G	A	G	A	G	A	G	A	975
Qy	336	Ser	Leu	Thr	Asn	Leu	Val	Phe	Val	Ser	Pro	Tyr	Met	Thr	Thr	Ala	355
Db	976	T	C	T	T	G	G	T	T	T	C	T	T	C	T	T	1035

Qy	356	Val	Val	Gly	Ile	Ser	Leu	Ala	Glu	Gly	Ile	Ala	Val	Gly	Arg	Ser	Phe	Ala	Met	Tyr	375
Db	1036	GT	CT	C	AGG	CA	TG	TT	GC	CT	TG	CA	CG	AA	GT	GT	AG	CA	AA	CA	1095
Qy	376	Lys	Asn	Tyr	Asn	Ile	Asp	Gly	Asn	Lys	Glu	Met	Ile	Ala	Ile	Gly	Thr	Met	Asn	Val	395
Db	1096	AA	AG	ACT	AC	CA	AA	TG	GT	TA	CA	AA	GA	GT	GT	AG	CA	AA	GT	AA	1155
Qy	396	Gly	Ser	Phe	Thr	Ser	Cys	Tyr	Leu	Thr	Thr	Gly	Pro	Phe	Ser	Arg	Ser	Ala	Val	Asn	415
Db	1156	GG	TT	CA	T	AC	CT	CT	T	GT	AT	GT	AT	CC	AC	GG	TT	CT	CT	CA	1215
Qy	416	Asn	Ala	Gly	Cys	Gly	Thr	Ala	Ala	Ser	Asn	Ile	Ile	Met	Ser	Leu	Ala	Val	Met	Leu	435
Db	1216	AT	GG	CG	GA	TG	T	CA	AA	CG	CA	GT	CT	CC	AA	CA	T	CA	T	CA	1275
Qy	436	Leu	Leu	Phe	Leu	Thr	Pro	Leu	Phe	His	Tyr	Thr	Pro	Leu	Val	Val	Leu	Ser	Ala	Ile	455
Db	1276	CT	T	CT	CT	CT	CA	C	CT	CT	TT	T	CA	AA	T	AC	CA	CC	AA	CA	1335
Qy	456	Val	Ser	Ala	Met	Leu	Gly	Leu	Ile	Asp	Tyr	Glu	Ala	Ile	His	Leu	Phe	Lys	Val	Asp	475
Db	1336	AT	CA	AG	CT	GT	CA	T	CT	TT	GG	T	CA	CG	TT	AA	T	CA	CT	CA	1395
Qy	476	Lys	Phe	Asp	Phe	Val	Val	Cys	Met	Ser	Ala	Tyr	Ile	Gly	Val	Val	Phe	Gly	Ser	Val	495
Db	1396	AA	GC	T	CA	TT	T	GT	T	GT	GG	G	CT	TT	T	T	T	T	T	T	1455
Qy	496	Ile	Gly	Leu	Val	Ile	Ala	Ile	Val	Ser	Val	Leu	Arg	Val	Leu	Phe	Ile	Ala	Arg	515	
Db	1456	AT	T	G	G	C	T	T	CT	TA	T	AG	CG	GT	AT	CT	TT	T	G	CA	1515
Qy	516	Pro	Arg	Thr	Phe	Val	Leu	Gly	Asn	Ile	Pro	Asn	Ser	Val	Ile	Tyr	Arg	Asn	Val	Glu	535
Db	1516	CT	T	AG	CA	C	AG	CA	T	CT	T	G	GA	AG	AT	AC	CG	GA	CT	CA	1575
Qy	536	Tyr	Gln	Asn	Ala	Lys	His	Val	Pro	Gly	Met	Leu	Ile	Leu	Glu	Ile	Asp	Ala	Pro	Ile	555
Db	1576	TAT	CT	CA	AG	CA	CT	AG	AT	T	CC	GG	AG	TT	T	T	G	CA	AT	T	1635
Qy	556	Phe	Ala	Asn	Ala	Ser	Tyr	Leu	Arg	Glu	Arg	Ile	Thr	Arg	Trp	Ile	Asp	Glu	Glu	Glu	575
Db	1636	TT	CT	CA	ACT	CC	CA	AT	TAT	GT	T	AG	GA	AG	GA	TT	CA	AG	AT	G	1695
Qy	576	Arg	Ile	Lys	Ala	Thr	Gly	Leu	Thr	Ser	Leu	Gln	Tyr	Val	Ile	Ile	Asp	Met	Ser	Ala	595
Db	1696	AT	G	T	G	GA	AG	CT	G	CA	AG	AT	T	GC	T	AG	GA	CT	T	CA	1755
Qy	596	Gly	Asn	Ile	Asp	Thr	Ser	Gly	Ile	Ser	Met	Leu	Glu	Val	Val	Lys	Ile	Thr	Glu	Arg	615
Db	1756	AC	G	GA	CA	CT	CA	T	AG	T	AG	T	TA	CA	AG	CA	CT	T	TA	AG	1815
Qy	616	Arg	Glu	Leu	Gln	Leu	Val	Leu	Val	Asn	Pro	Val	Ser	Glu	Val	Met	Lys	Lys	Leu	Asn	635
Db	1816	CG	A	CA	CT	CA	T	AG	TT	CT	T	AG	CG	AA	T	CA	CG	CG	T	CA	1875
Qy	636	Ser	Lys	Phe	Gln	Asn	His	Ile	Ser	Gly	Lys	Trp	Ile	Tyr	Leu	Thr	Val	Glu	Ala	Val	655
Db	1876	T	C	T	CA	CT	T	T	G	CG	CA	CT	T	T	AG	CA	CA	CA	AA	AA	1935
Qy	656	Gly	Ala	Cys	Asn	659															
Db	1936	GAT	T	CT	T	G	CT	C	1947												
RESULT 14																					
ADA70110																					
ID	ADA70110 standard; DNA; 2000 BP.																				
XX																					
AC	ADA70110;																				
XX																					
DT	20-NOV-2003 (first entry)																				
XX																					

RESULT 14

ADA70110

ID ADA70110 standard; DNA; 2000 BP.

XX

AC ADA70110;

XX

DT 20-NOV-2003 (first entry)

XX

Db 1486 GCCTTGTCAGTTGTCAGGGCGTTGCTGTACGTGGCCAGCCCTGCCACCTGCAAGCTCGG 1545

QY 523 AsnIleProAsnSerValIleTyrArgAsnValGluHIStyrGlnAsnAlaLysHisVal 542

Db 1546 AACATAGCAGGAGCGAGACCTTCGCGCAGCTGAAGCAATACCCCAAGCGAAGCATC 1605

QY 543 ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrIleu 562

Db 1606 CTTGGCATCTTCGCTTCGACATCGGCTCCCATCTACTTCGTCATGCGGGTTACTCTG 1665

QY 563 ArgGluArgIleThrArgTrrPileAspGluGluGluArgIleLysAlaThrGlyGlu 582

Db 1666 CGAGAAAGGATTTGAGATGGGTGAAGATGAGGATAACCTGTCAAGAGCGTCGGGCAC 1725

QY 583 ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly 602

Db 1726 ---GATCTGCAATACTTGGTTCTTGATCTTGGTGTGTCACTTCTGTGCAAACTCTGGA 1782

QY 603 IleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLeu 622

Db 1783 GTTGGGATGCTACTAGAAAGTACACAAGAGCCTCGAAGAGGATCAGCATAGCTCTG 1842

QY 623 ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641

Db 1843 ACGAATCCGAGGCTAGAGTTAACAGAGAGCTGGTGTCTGTGATAGTCAGGAGCATC 1902

QY 642 LeuGlyLysLysTrrPileTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661

Db 1903 TTAGGGATGAGTGGGTCTTCTGACGCTCAAGACGCCATCAGCGGCTGCGATACGCG 1962

QY 662 LeuArgAlaSerLysThrAsnProLysLysAspGlu 673

Db 1963 CTGCAGATATCCAGA---AATAAGGAGAGACGAA 1995

RESULT 15

AAZ50490

ID AAZ50490 standard; cDNA; 2311 BP.

XX

AC AAZ50490;

XX

DT 23-MAY-2000 (first entry)

XX

DE Wheat sulphate permease-1 cDNA clone.

XX

KW Sulphate Permease; sulphate assimilation protein; wheat; probe; mapping;

KW marker; plant breeding; chimeric gene; transgenic plant; antibody;

KW screen; ss.

XX

OS Triticum aestivum.

XX

PH Key Location/Qualifiers

FT CDS 100..2070

FT /*tag= a

FT /product= "Wheat sulphate permease-1"

FT /note= "Derived from clone wlk1.pk0028.el"

XX

PN W0200004154-A2.

XX

XX

PD 27-JAN-2000.

XX

PP 13-JUL-1999; 99WO-US015810.

XX

PR 14-JUL-1998; 98US-0092833P.

XX

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX

XX Allen SM, Falco SC, Thorpe CJ;

XX

XX WPI; 2000-195025/17.

DR P-PSDB; AAY44944.

XX

XX Nucleic acid fragments encoding sulfate assimilation proteins in plants

PT and seeds useful as probes for isolating cDNAs and genes encoding

PT homologous proteins, in producing transgenic plants.

XX Claim 3; Page 55-56; 79pp; English.

XX The present sequence is a cDNA clone encoding wheat sulphate permease, a
 CC sulphate assimilation protein. This sequence is obtained from wlk1
 CC library, clone wlk1.pk0028.el, derived from wheat seedlings, 1 hour after
 CC treatment with fungicide. This sequence is used as a probe to isolate
 CC other plant sulphate assimilation proteins, for genetic and physical
 CC mapping of related genes and as markers of traits linked to the gene.
 CC This is useful for plant breeding and to construct chimeric genes, used
 CC to create transgenic plants with altered levels of sulphate permease. The
 CC sulphate permease peptides are useful for producing antibodies, that are
 CC used to screen and isolate cDNA clones

SQ Sequence 2311 BP; 670 A; 493 C; 537 G; 611 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,73e-175 Length: 2311
 Score: 1799.00 Matches: 336
 Percent Similarity: 72.30% Conservative: 126
 Best Local Similarity: 52.58% Mismatches: 172
 Query Match: 51.77% Indels: 5
 DB: 3 Gaps: 2

US-10-762-049-18 (1-680) x AAZ50490 (1-2311)

QY 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60

Db 187 GTGTACAAAGTGGGCTATCCCCCTCGAAGAACTTGGCCACAGATTACAGAAACATTG 246

QY 61 LysGluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80

Db 247 AGGGAGACTTCTTCCACGACAAACCCGCTGCGTCAGTATAAGGCCCAATCCGACCCAGG 306

QY 81 LysPheMetLeuGlyLeuGlnPhePheProIlePhePheLeuTrpAlaProLysTyrThr 100

Db 307 AGGTTTCATGATGGGCTGGAGTTCTTGTTCCTATATTTGGGTGGGTAGGAGTTACAGT 366

QY 101 PheGlnPheLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120

Db 367 CTCACAAAGTTCAAAGCGCATCTGATTGCGGATTGACCATCGCAAGTCTCTGTATCTCT 426

QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140

Db 427 CAGGACATTGGCTATTTCGAAGCTTGCTAATCTGGATCCGAGTATGGGCTTTACTCCAGC 486

QY 141 PheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160

Db 487 TTCATTCTCCATTGATCTATGCTGCAATGGGTAGCTCAAGGGATATAGCATGTGTCCTCA 546

QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180

Db 547 GTTGCTGTGGTTCTCTTTTGTAGGTTTCTTCTACAGCTGAGGTTGACCATGTCAAA 606

QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200

Db 607 AACAAAGGAGGAATACATCGCCTCGCTTTTCAGGCAACCTTCTTCGCTGTATCACTCAA 666

QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220

Db 667 GCAGCCTTAGGATTTCTTAAGGTTAGGATTCCTTATAGATTCTTGTGCGATGCTGGGATT 726

QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeu 240

Db 727 GTGCGATTTCATGGGGAGCTGCCATTACTATTGCCCTGCAGCAGCTGAATACAGTGTTC 786

QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260

Db 787 GGCATCGCAAACTTTTACAAAGGAAACCGACATAGTTTCTGTCTCATGGAATCTGTCTCGAGA 846

QY 261 GlnThrHisGlu---TrpArgTrpGluSerAlaValLeuGlyCysValPhePhePhe 279

847	Db	TCAGTTTCATCATCAGCGGTGGAACCTGGCAGACAAATGTTGTATGTTGGCGGTATCTTTCCTGGTTTTC	906
280	Qy	LeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheTyrPheTyrValSerAlaMet	299
907	Db	CTTCTGTTGCGAAGTACATCGGAAGAGNAGAAAGGAAGCTTTTCTGGGTGCCAGCTATT	966
300	Qy	AlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLys	319
967	Db	GCTCCTATAATTTTCAGTGATTTCTAGCAACATTTTTTGTATACATTACTCGTGGCGACAAG	1026
320	Qy	HisGlyValGluValIleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsn	339
1027	Db	CAAGGAGTTTCAGATAGTGAAGCACATGAACAGAGGAATCAACCCATCATCAGTACACAAG	1086
340	Qy	LeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIle	359
1087	Db	ATTATTATTCACGGCCCATTTGTTGCAAAAGGTTTCAAGATCGGTGTTGTTTGGCGGATA	1146
360	Qy	IleSerLeuAlaGluGlyIleAlaValIleGlyArgSerPheAlaMetTyrLysAsnTyrAsn	379
1147	Db	GTTGGTTTGACAGAAGCTGTAGCTATTGGAAGGACATTGTCGTCTATGAAGGACTACCAG	1206
380	Qy	IleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThr	399
1207	Db	TTAGATGGAACAAGAGAGATGGTAGCATTTGGAACCATGAACATAGTAGGCTCAATGACA	1266
400	Qy	SerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCys	419
1267	Db	TCYTGCTATGTACACACAGGTTCTTTCTCAGTTTCGGCAGTTAACTTCATGGCTGGCTGC	1326
420	Qy	LysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeu	439
1327	Db	AAGACTCCTGTATCCATGTTGTTATGTACAGTAGTGGTTCTCTTACCTTGTGGTCATC	1386
440	Qy	ThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMet	459
1387	Db	ACACCGCTATTCAAATATACACCGAAATGCATCTCCTAGGTCGATCATTAATTTCTGCGGTG	1446
460	Qy	LeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPhe	479
1447	Db	ATCGGCTTTGTGGACTACGAAGCAGCAATTTCTCATCTGGAAGTTGCAAAATTTGGACTTC	1506
480	Qy	ValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVal	499
1507	Db	ATTGCTTGTCATGGGAGCTTTTTCCTGGGTGTTGTTTGTATCCGTTGAGATTGGCTCTTG	1566
500	Qy	IleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPhe	519
1567	Db	ATTGCTCTAGCAATCTCATTTGGCCAAATATCTTCTTCAAGTAAACAGGCCACAGACAGCC	1626
520	Qy	ValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAla	539
1627	Db	CTACTTGGAAACCTTCCCGGCACCCTATATACCGGAACATCAGCCAGTATCCAGAAGCA	1686
540	Qy	LysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAla	559
1687	Db	AAACTTACTCTCGGGTGGTCATTTGTGAGGGTTGATTCTGCTATTATTTTCCACTCT	1746
560	Qy	SerTyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAla	579
1747	Db	AAATTACGCGAAGAAAGAAATCTTAGTGGTGGCTGACACAGCAAGCAAGACAGAGCTAAAGCA	1806
580	Qy	ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp	599
1807	Db	GTGGGATTGCTTAAATTCAGTTTCTGATTGTGGAAATGTGCGGGTCATCGCATCATGAT	1866
600	Qy	ThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGln	619
1867	Db	ACAAGCGCATACATGCTCTTGAAGATCTATACNAGAACTCTTCAGAAAAAAGATATGCGAG	1926
620	Qy	LeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln	639
1927	Db	CTCAATCTGTGCAATCTGGTTCGTGCTCATAGAAAAATCGCAAGGCTCGAAGCTCACC	1986

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 31, 2005, 11:52:46 ; Search time 331 Seconds
(without alignments)
3361.535 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNTI.....NLRAKTNPKUDETEGNNV 680

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spo1/US10762049/runat_30082005_121246_6006/app_query.fasta_1.839
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10762049 @CGN 1 1 69 @runat_30082005_121246_6006 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3475	100.0	2449	4	US-09-720-317A-17
2	2164	62.3	1981	4	US-09-720-317A-3
3	1901	54.7	2067	4	US-09-720-317A-15
4	1844.5	53.1	2279	4	US-09-720-317A-1
5	1799	51.8	2311	4	US-09-720-317A-19
6	963.5	27.7	2022	4	US-09-720-317A-21
7	731.5	21.1	780	4	US-09-720-317A-7
8	687.5	19.8	1791	4	US-09-252-991A-7241
9	676.5	19.5	1767	4	US-09-252-991A-7184
10	675.5	19.4	1240	4	US-09-720-317A-5
11	651	18.7	2441	4	US-09-785-381-4
12	650	18.7	4113	4	US-09-785-381-2

13	643.5	18.5	1773	4	US-09-902-540-7746	Sequence 7746, Ap
14	643.5	18.5	4854	4	US-09-902-540-768	Sequence 768, App
15	641.5	18.5	2832	4	US-09-949-016-4516	Sequence 4516, Ap
C 16	625	18.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 17	619	17.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
18	612	17.6	1818	4	US-09-602-787A-615	Sequence 615, App
19	586.5	16.9	2858	4	US-09-949-016-5349	Sequence 5349, Ap
20	586.5	16.9	2862	1	US-08-424-567-1	Sequence 1, Appli
21	586.5	16.9	2882	2	US-08-711-928-1	Sequence 1, Appli
22	586.5	16.9	2882	3	US-09-184-937-1	Sequence 22, App
23	578	16.6	2682	4	US-09-614-221A-222	Sequence 5627, Ap
24	577	16.6	4927	4	US-09-949-016-5627	Sequence 25, Appl
C 25	571.5	16.4	31147	4	US-09-596-002-25	Sequence 6643, Ap
26	567	16.3	2487	4	US-09-248-796A-6643	Sequence 5, Appli
27	548.5	15.8	1992	4	US-09-875-811-5	Sequence 9, Appli
28	548	15.8	1971	4	US-09-875-811-9	Sequence 13, Appl
29	548	15.8	2420	4	US-09-875-811-13	Sequence 1, Appli
30	545.5	15.7	2040	4	US-09-875-811-1	Sequence 249, App
31	489	14.1	1851	4	US-09-543-681A-249	Sequence 7, Appli
32	485.5	14.0	1818	4	US-09-875-811-7	Sequence 11, Appl
33	485	13.9	1797	4	US-09-875-811-3	Sequence 3, Appli
34	482.5	13.9	1866	4	US-09-875-811-3	Sequence 16258, A
35	464	13.4	8774	4	US-09-949-016-16258	Sequence 6, Appli
36	451	13.0	2913	4	US-09-795-927-6	Sequence 8, Appli
37	451	13.0	3749	4	US-09-795-927-8	Sequence 5234, Ap
38	449	12.9	1608	4	US-09-252-991A-5234	Sequence 5199, Ap
39	449	12.9	1911	4	US-09-252-991A-5199	Sequence 6127, Ap
40	426.5	12.3	1419	4	US-09-902-540-6127	Sequence 286, App
C 41	426.5	12.3	2070	4	US-09-902-540-286	Sequence 1, Appli
C 42	402.5	11.6	1230230	4	US-09-438-185A-1	Sequence 1, Appli
C 43	389.5	11.2	1230025	4	US-09-198-452A-1	Sequence 7460, Ap
C 44	372	10.7	1116	4	US-09-252-991A-7460	Sequence 372, App
45	369	10.6	1701	4	US-09-543-681A-372	

ALIGNMENTS

RESULT 1
US-09-720-317A-17
; Sequence 17, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 2449
; TYPE: DNA
; ORGANISM: Glycine max
US-09-720-317A-17

Alignment Scores:
Pred. No.: 0 Length: 2449
Score: 3475.00 Matches: 680
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-762-049-18 (1-680) x US-09-720-317A-17 (1-2449)

Qy 1 HisGluLeuAlaArgThrLeuSerTyrIleThrHisfileCysLeuLeuArgAenThrIle 20
Db 2 CACGACTAGCTCCACATTAAGTTATATACACATATTGCTTGCTTAGAATATTATT 61

; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-3

Alignment Scores:

Pred. No.: 2,17e-258 Length: 1981
Score: 2164.00 Matches: 403
Percent Similarity: 86.09% Conservative: 86
Best Local Similarity: 70.95% Mismatches: 79
Query Match: 62.27% Indels: 0
DB: 4 Gaps: 0

US-10-762-049-18 (1-680) x US-09-720-317A-3 (1-1981)

```
QY 105 LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer 124
DB 8 GAGTCCGACCTGATCGCGCGCATCACCATCGCCAGCCTCGCCATCCGCGCAGGGGCATCAGC 67

QY 125 TyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIleProPro 144
DB 68 TAGCCCAAGCTCGCCCAACCTGCCGCCCGCTGCTCGGACTCTACTCGAGCTTCGTGCCGCCG 127

QY 145 LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly 164
DB 128 CTGGTGATACGCTGATGGGAGCTCCCAAGACCTGGCGGTGGGACCGTGGCGGTGGCG 187

QY 165 SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeu 184
DB 188 TCGCTGCTCATAGCTCCATGCTCGGCAGCGAGGTGTCGCCGAGCGGAGAACCCCGTGCCTC 247

QY 185 TyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGly 204
DB 248 TACTGTCACTCGCCTTCACGCCACCTTCTTCGCCGGCGTCTTCAGGGCTGCTCGCCG 307

QY 205 LeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMet 224
DB 308 CTCCTCAGGTTGGGCTTCATCGTGACCTGCTGTCGCACGCGACGATCGTGGGGTTTCATG 367

QY 225 GlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis 244
DB 368 GCCGGCGCGCGCAGCGGTGTGCTGTCAGCAGCTGAAGGCGCATGCTGGGCCCTCGTCCAC 427

QY 245 PheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlu 264
DB 428 TTCAACCACTCCACCGACGCTGCTCGCTCATGGATCCGCTTCACGCCACACACACAG 487

QY 265 TrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArg 284
DB 488 TGGCGGTGGGAGACGCTGCTCGGCTGCGGCTTCCTCTTCCTCTCTCTCTCTCTCTCTCT 547

QY 285 TyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSer 304
DB 548 TTCAATCAGCAAGAGCGCTCCACAGCTGTTTGGATCTCCGCGCGCGCGCTTGGACGTCC 607

QY 305 ValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluVal 324
DB 608 GTCGTGTCCGGAGCGTCTGTTGTATCTCAGCAGCAGCTGAAACACACCGCATCGAAGTG 667

QY 325 IleGlyLeuLeuLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSer 344
DB 668 ATCGGTTTACCTGAAGAAGGCTGAATCCACCGTCGGTGACAAAGCCTGCAATTCTCACCG 727

QY 345 ProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGlu 364
DB 728 CCCTTACATGATGCTCGCGCTCAAGACTGGGATCATCCCGCGTCAATTGCCCTCGCCGAA 787

QY 365 GlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLys 384
DB 788 GGAATCCCGTGGGGAGGAGCTTCGCCATGTTCAAGAACTACACATGACGGACCAACAG 847
```

```
QY 385 GluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThr 404
DB 848 GAGATGATCGCGATCGGACGATGAACGCTCGTGGCTCGCTACGCTGCTACCTGACC 907

QY 405 ThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSer 424
DB 908 ACGGGGCGCTTCTCGCGCTCCGCGCTGAATACACGCCGGTGCAGGCGCATGTCG 967

QY 425 AsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHis 444
DB 968 AACGTGGTCACTGTCGCTGGCGGTGATGTCACGCTGCTGTCGACGCCGCTGTTCCAC 1027

QY 445 TyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAsp 464
DB 1028 TACACGCCGCTGCTGCTGTCGCGCATCATGCTCCGCGATGCTGGGCGCTGGTCGAC 1087

QY 465 TyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSer 484
DB 1088 TTCGGGCGCGCTGACCTGTCGCGCTGTCGACGATGTCGCTGCTGCTGCTGCTGCTGCTG 1147

QY 485 AlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIle 504
DB 1148 GCGTACCTGGCGCTGCTTTCGGCAGCGTTCGAGGTGCGCTGCTGCTGCTGCTGCTGCTG 1207

QY 505 SerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIle 524
DB 1208 TCCCTGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1267

QY 525 ProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGly 544
DB 1268 CCCGCCACCATGCTGACCGGAGTGGACGATGTCGCCCGCGCGACGCGTGCCTGCTGCTG 1327

QY 545 MetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGlu 564
DB 1328 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1387

QY 565 ArgIleThrArgTrpIleAspGluGluGluGluGluGluGluGluGluGluGluGluGlu 584
DB 1388 AGGATCTCGCGTGGATCGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1447

QY 585 LeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSer 604
DB 1448 GTCCGTCAGTGTCTCTGACATGGGTGCTTGGACGAGGAGGAGGAGGAGGAGGAGGAGG 1507

QY 605 MetLeuGluGluValLysLysIleThrGluArgGluLeuGlnLeuValLeuValAsn 624
DB 1508 ATGCTGGACGAGCTCAACAGTCTTGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1567

QY 625 ProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLys 644
DB 1568 CCGCGCAGCAGCATCATGAAGAAGCTGACAGCTCCAAAGGTGCTGGAGCAGATCGGCGAC 1627

QY 645 LysTrpIleTyrLeuThrValGluAlaValGlyAlaCysAsnPheAsnLeuArgAla 664
DB 1628 GAGTGGGTGTCCCGAGGTTGGCGGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1687

QY 665 SerLysThrAsnProLysLysAsp 672
DB 1688 CACAAGCCGGGAATGGCCAAAGGAC 1711
```

RESULT 3

US-09-720-317A-15
; Sequence 15, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21


```

; Sequence 1, Application US/09720317A
; Patent No. 696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR FILING DATE: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2279
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-1

Alignment Scores:
Pred. No.: 1,69e-218 Length: 2279
Score: 1844.50 Matches: 355
Percent Similarity: 74.25% Conservative: 115
Best Local Similarity: 56.08% Mismatches: 158
Query Match: 53.08% Indels: 5
DB: 4 Gaps: 3

US-10-762-049-18 (1-680) x US-09-720-317A-1 (1-2279)

Qy 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
189 GTGCACAAAGGTGGCGCGCCGACCGCGCGAGCAGCGCGAGATGAAGGTGAGGGTG 248
Qy 61 LysGlnThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 AAGGAGACCTTCTTCCCGAGCAGCCCGTTCGGGCGGTTCAGGGGCGAGCCCGGGGACG 308
Qy 81 LysPheMetLeuGlyLeuGlnPhePhePheProilePheGluTrpAlaProLysTyrThr 100
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
309 CAGTGGCTCATGGCGGTGAGTACCTCTTCCCATCTGAGCTGGGTGGGTGAGTACTCC 368
Qy 101 PheGlnPheLysAlaAspLeuLeuAlaGlyLeuThrLeuAlaSerLeuAlaLeuPro 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 TTGTGCTCTTCAAGTCCGACCTGTCGCGGCGCTCACCATGCCAGCCTCGCATTCCT 428
Qy 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnProProileLeuGlyLeuTyrSerSer 140
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 CAGGGCATTAAGTACGCGAAGCTGGCAAGCTTGCTCCCATATTCGGGCTGTATTCGAGC 488
Qy 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
489 TTGTGCGCGCGATGGTGTACGCGGTCTGGGAGCTCCCGTGACCTGGCGGTGGGCGCG 548
Qy 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
549 GTGTGCATCTCGTGTGATCATGGGTGCTCATGCTGGCGAGCGCGCTGAGCCCATCGCG 608
Qy 181 AspProLysLeuTyrLeuHisAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
609 GAGCGGAGCTGTTCCTGAGCTGGCCCTTACCTCCACCTGTCGCGGCGGTGGTGCAG 668
Qy 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
669 GCCTCCCTGGGCATCTCGGCTGGCTTCGTTCATCGACTTCCTGTCCAAAGGCGAGCGCTG 728
Qy 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLysSerIleLeu 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
729 GTGGGTTCATGGCGCGCGCGCCATCATGTGGCGCTGAGCAACTCAAGGGGCGTCTG 788
Qy 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
789 GGCATCTGCCACTTCACCAACCGAGATGGGCATGTCCTCCAGTCATGGCCTCCGCTTCCAC 848

```

```

261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
849 CACACCAGCAGTGGTCTGGCAGACGATCCTCATGGCGCTGCTTCTCTGCTTCTCTG 908
Qy 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
909 CTGTGGCGGAGCATGTGAGCATCAGATGGCCAAAGCTTTTCTGGGTTCGGGTGGCGG 968
Qy 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
969 CCCCTGGCATCGGTCAACCATCTCGACGCTGCTTGTCTTCTCTTCAAAGCTCAGAACC 1028
Qy 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1029 GGCATCAGCATCATTTGGCAGCTCAAGTGGCGCTGAATCGCCCTCGTGGGCAAGCTC 1088
Qy 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1089 CTGTTTGACACAGCGGTATTTAGGCGCTCACCATGAAGACTGGCCTTGTACCGGAATCATC 1148
Qy 361 SerLeuAlaGluGlyIleAlaValArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1149 TCACTCAGCGAAGGATAGCGGTGGTAGACATTTGCTCACTCAAGGACTACCAAGATA 1208
Qy 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1209 GATGGAAACAAGGAGATGATGGCCATAGGGTTGATGAATGTTGTTGGTCTCTGCACATCA 1268
Qy 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1269 TGTACTGTAACAACAGGTGGTCTCCCGCTCTGCTGTAACACCAACGCGCGCTGCAAG 1328
Qy 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1329 ACTGCCATGTCCACGCTGATCATGGCGCTGACTGTGATGGTCAAGCTGCTTCTCATG 1388
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1389 CCACCTGTCGTGTACACACCAACGTTGCTCGGAGCGATCATCATCGCGCGGTGATC 1448
Qy 461 GlyLeuIleAspTyrGluAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1449 GGCCTCATGATTTCCCGCGGTGGTACCACATCTGGAAGATGGACAGATGGATTTTCTG 1508
Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1509 GTGTGGTGTGGCGGTTCGCGCGCTCATCTTCATCTCAGTCCAGAGGCGCTTGGGATA 1568
Qy 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1569 GCGGTTGGTATATCTATATTAGGGTGTGTGATGCAGATCACAGGCGCAAGATGATGTT 1628
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1629 CAAGGAAACATCAAGGGGACTGATATTTACAGAGACTGCATCCTACACAGGAGGCCAA 1688
Qy 541 HisValProGlyMetLeuIleLeuAspAlaProIleTyrPheAlaAsnAlaSer 560
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1689 AGAGTTCTTGGGTCTTGTATCTTGGCCATTTGAAGCACCGGATAAATTCGCCCACTCCA 1748
Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1749 TACCTGAATGAAGGATTAAGATGGATA-----GAGGAAGATCTTTTGAACAGGAT 1802
Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1803 AAACATACATACTCCATTTCATTAATCTTGGATCTGTGAGCTGCTCCTGCAATTCACACA 1862
Qy 601 SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1863 AGTGGCATAGCGTCTCTCATTTGACATTAAGAAATCAATAGAGAAACGTGCTGGAGCTT 1922

```



```

Db 1687 AAATTACTCTCTGGGGTGGTGAATGTGAGGGTGAATCTCTGCTATTATTTTTCACACTCT 1746
Qy 560 SerTyrLeuArgGluArgIleThrArgTrrPileAspGluGluGluArgIleLysAla 579
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1747 AATTACGCTCCGAGAAAGAAATCTTAGGTGGCTGACAGACGAAAGACAGACAGCTAAAGCA 1806
Qy 580 ThrGlyCluThrSerLeuGlnTyrValIleLeuAspMetSerAlaValGlyAlaMet 599
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1807 GTGGGATTTGGCTAAATCAATTCCTCGATTTGTGAATGTGCGGGTTCATCGACATCGAT 1866
Qy 600 ThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGln 619
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1867 ACAAGCGGCATACATGCTCTTGAAGATCTATACAGAACTTTCAGAAAAAGATATGCAG 1926
Qy 620 LeuValLeuValLeuProValSerGluValMetLysLysLeuAsnLysSerLysPheGln 639
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1927 CTCAATCTGTGGAATCTGCTGCTCGCTCATAGAAAACTGCAAGCGTCAAGCTCACC 1986
Qy 640 AsnHisLeuGlyLysLysTrrPileTyrLeuThrValCluGluAlaValGlyAlaCysAsn 659
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1987 GAGCACATTTGGAAGCAATATATCTCTCGCGTCTCTGACGCTGTGCGAATTCGTACG 2046
Qy 660 PheAsnLeuArgAlaSerLysThrAsnPro-LysLysAspGluThrGluGlyTrrP 677
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2047 -----ACGAAGTCGATGCAGAACCTGTGACGCAAGTAGTTCGGAGGAATGG 2092

RESULT 6
US-09-720-317A-21
; Sequence 21, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-720-317A-21

Alignment Scores:
Pred. No.: 6,81e-109 Length: 2022
Score: 963.50 Matches: 206
Percent Similarity: 56.30% Conservative: 129
Best Local Similarity: 34.62% Mismatches: 243
Query Match: 27.73% Indels: 17
DB: 4 Gaps: 7

US-10-762-049-18 (1-680) x US-09-720-317A-21 (1-2022)
Qy 91 ProilePheGluThrAlaProLysTyrThrPhe---GlnPheLeuLysAlaAspLeuIle 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 CCGTGGCTGGCGTGGATGGCGAGCTACAGATGGAAGAGACTTCACAGCCGACCTCGCC 125
Qy 110 AlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAla 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 126 GCCGGCATCATCTGCGCGCTATGCTTGCTCTGAGCAATGTCATATGCAAGCTGGCT 185
Qy 130 AsnLeuProLysLeuGlyLeuTyrSerPheIleProLysLeuTyrAlaMet 149
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 186 GGGCTTCACCAATTTATGGGCTCTACACAGGCTTTGTCCCACTATTATTGTCTACGGGAT 245
Qy 150 MetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuMetGly 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 246 TTTGGGCTCTCACACAAATAGCAGTAGGTCCAGTGGCACTTGTCTCTGCTAGTGTCC 305

```

```

Qy 170 SerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAla 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 AATGTTCTTTGGGGTAGTGT-----AATTCATCTAGTGAGCTGTACACGGAATTAGCC 359
Qy 190 PheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGly 209
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 ATATATTGGCATTCATGGTTGGAATATCGGAATGCTTGTGATGCATTTCTTAAGACTTGGC 419
Qy 210 LeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThr 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 TGGCTTATTCGTTTTCATTAGCCATTCCTGTAATATCTGGATTCTACTACAGCTTCGGCCATC 479
Qy 230 ValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAla 249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 480 GTAATTTGGTTTGTCCCAATCAAGTATTCTTGGGTACAGT---GTTACAAGAAGTAGC 536
Qy 250 AspIleLeuSerValMetArgSerValPheThrGlnThrHisGluThrArgTrrPrrGluSer 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 537 AAAATTTATACCTTATTTAGAGATATAATGCTGGAATAGATCAGTTCTCTCGCCCTCCA 596
Qy 270 AlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSerLysLys 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 597 TTTGTAATGGGATCAGCGTTTCTTGTATTCTTCTAATAATGAAGAGCTAGGGAACA 656
Qy 290 ArgProArgPhePheTrrPrrValSerAlaMetAlaProLeuThrSerValIleLeuGlySer 309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 657 AATAAAAAAATTACGTTTCTCTGAGCTTCTGTGCTCCACTAACAGCTGTTGTTCTTGAACA 716
Qy 310 LeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLys 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 717 TTGTTTGTGAAAAATTTTCCGTCCAACT-----GCCATATCAGTGTGTGAATATACCG 770
Qy 330 LysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerPro----- 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 771 CAAGGCTT-----CCCACT-----TTCTCCATTTCTCGAGGATTTGAA 809
Qy 346 TyrMetThrAlaValLysThrGlyIleValGlyIleValGlyIleSerLeuAlaGluGly 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 810 CATCTGATCCCTAATGCTCAACTGCAATACTTATCACTGGTGTGCTATTGTTGGAGTCT 869
Qy 366 IleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGlu 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 870 GTTGGGATTTGTAAGCGTTAGCTGCGAAGATGGTTAGTGTGAGTTGACCTCAACAAGAG 929
Qy 386 MetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThr 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 930 TTATTTGGCTTTGGCTTATCAATATATATCGGGTTCAATTTCTCTGCATATCTCTGTACA 989
Qy 406 GlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsn 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 990 GGCCTCTTTCTAGGTCTGCTGTAATCATGAAAGCGGGCAAGACTGGATTATCAGGA 1049
Qy 426 IleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyr 445
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1050 ATCATAATGGGCATATAATAATTTGCAGTGTCTCTCTGTTTATGACACCATTTATTACTGAT 1109
Qy 446 ThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyr 465
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1110 ATACCTCAGTGTGCAATTTGGCTGCCATTGATTTCTGCTGTCTGCTGCTGTGATATTAT 1169
Qy 466 GluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAla 485
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1170 GAAGAGGCCATCTCTCTGTGGGTATGTAAGAAGGATTTCTTCTGTGGGGCGATGACA 1229
Qy 486 TyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSer 505
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1230 TTTTACTACAACTTAATTTTGGCATTGAGTTGGTGTCTCTGTTGGGGTGGGTTTTCG 1289
Qy 506 ValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIlePro 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1290 CTGGCATTGTGATCCATGAATCTGCNAATCGCATATAGCTGTGTTTGGGCGGTTGCCT 1349

```

Qy	526	AnSerValIleTyrArgAsnValgluHisTyrGlnAsnAlaLysHisValProGlyMet	545
Db	1350	GGCACCACTGTGTACAGAAATACATTTGCAGTACCCTGAGGCTTATACATACACACGGATT	1409
Qy	546	LeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArg	565
Db	1410	GTTGTGTGTCGTGTGTATGCACCAATCTACTTTCTAACATPAAGTTTACATAAAGGACAGG	1469
Qy	566	IleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeu	585
Db	1470	TTGCGTGAGTATGAGCTCAAACTCCCAATTCAAACCGTGCACCTGATTTGGNAGGGTG	1529
Qy	586	GlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMet	605
Db	1530	TACTTTGTGATCCCTCGAGATGTCCTGTTACATACATTCGACTCGAGCGCTGTTCAAGCT	1589
Qy	606	LeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnPro	625
Db	1590	CTCAAGGACCTGCACCAAGAAATACAAGACGCGCACATCCAGATTTGTTATAGCAATCCT	1649
Qy	626	ValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLys	645
Db	1650	AACCGGAGGTGCACCTATTCTCTCAAGACGGCGCATCTCGACATGATTTGGCGCAGGG	1709
Qy	646	TrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSer	665
Db	1710	TGGTGTTCGTCGCGAGTGCACGACGCGGTGCAAGTATGCTTCCAGCATGTGCGGAGTTCA	1769
Qy	666	LysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal	680
Db	1770	TCGTGCAATGCCATTAAGTTATCCCCACAGCGCTCTGGGAACCTTG	1814
RESULT 7			
US-09-720-317A-7			
; Sequence 7, Application US/09720317A			
; Patent No. 6696292			
; GENERAL INFORMATION:			
; APPLICANT: Stephen M. Allen			
; APPLICANT: Saverio C. Falco			
; APPLICANT: Catherine J. Thorpe			
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins			
; FILE REFERENCE: BB-1167			
; CURRENT APPLICATION NUMBER: US/09/720,317A			
; CURRENT FILING DATE: 2000-12-21			
; PRIOR APPLICATION NUMBER: 60/092,833			
; PRIOR FILING DATE: 14-07-1998			
; NUMBER OF SEQ ID NOS: 31			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 7			
; LENGTH: 780			
; TYPE: DNA			
; ORGANISM: Helianthus tuberosus			
US-09-720-317A-7			
Alignment Scores:			
Pred. No.: 9.44e-81 Length: 780			
Score: 731.50 Matches: 145			
Percent Similarity: 79.64% Conservative: 31			
Best Local Similarity: 65.61% Mismatches: 38			
Query Match: 21.05% Indels: 7			
DB: 4 Gaps: 1			
US-10-762-049-18 (1-680) x US-09-720-317A-7 (1-780)			
Qy	467	AlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyr	486
Db	17	GCCCGCATTCACCTCTCGACACTAGACAAATTCGACTTTGTTGATGCATGAGTGCTAC	76
Qy	487	IleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerVal	506
Db	77	TTTCGTGTGTCTTTGGAGTGTGAATTTGATTTATCGCGGTCGCATGTCGTTG	136
Qy	507	LeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGluValAsnIleProAsn	526

Db	73	CCACTCGCCCGCTGGGTACCCGGAGCTCGATAGCTGCTCTCACTACCGCGCGCGCTGTTCT	132
Qy	105	LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer	124
Db	133	CGTCCGGATGTCAGGCGGGGCTCGGTAGCGCAATCCAGATCCCAACCGCCATCGCT	192
Qy	125	TyrAlaIysLeuAlaAsnLeuProProIleLeuGlyIleuTyrSerSerPheIlePro	144
Db	193	TACGCCCAGATCGCGGCTTCCCGCCAGGTAGGAGCTCTATGGTGTGATCTCTCCGATG	252
Qy	145	LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyIleThrValAlaValGly	164
Db	253	CTGATCTACGCCCTTGATCGGCAGCTCGCGCAGTTGATGTAGGCCCGCAGCGCGCG	309
Qy	165	SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspPro-----AsnGluAspPro	182
Db	310	-----ACCGCGCGCATGTCGCGCGCGCCATCAACCCGCTGGCGGGCGGATCCG	360
Qy	183	LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla	202
Db	361	CAGCGCTGTGTGGACCTGTGATGATCGTCGATCATGTGTCGGCTGTTCTCCATCGTC	420
Qy	203	LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGly	222
Db	421	GCCGGCTTGGCCCGCGCGGTTTCATCGCCAGCTTCTCTCGCAGCATCTCGTGGCG	480
Qy	223	PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuIysSerIleLeuGlyLeu	242
Db	481	TACCTCAACGCATCGGCCTGAGCCTGCTGGTCGGGCNACTGGGCAGCTGTTCTGGTAC	540
Qy	243	GluHisPheThrHisGly-----AlaAspIleIleSerValMetArgSerValPheThr	260
Db	541	GAGCGCGCACCGCGGCTTCGTCGCCGCGATCTCTCGCGTCTGTGGAGAACCTGCTG	597
Qy	261	GlnThrHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeu	280
Db	598	-----CACATCCATTGCCGACGCTGATCTCTCGCAGCCTCTCGCTTGTCTGATGA	648
Qy	281	LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTyrValSerAlaMetAla	300
Db	649	GTGCTGTG-----CCGCGCGCGCTTC-----CCGAGTTGCCCGCGG	684
Qy	301	ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis	320
Db	685	CGCTCTCGCGCGTCTCTGCGCAGCCTCGCGCGCGTCTGTCGGCTGGATCGCTAC	744
Qy	321	GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeu-----	337
Db	745	GGCGTGAATGCTGTGCGAGGTACCGCGCGGCGCTG-----CCGCAACTGAGCTGGCCG	798
Qy	338	---ThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleVal	356
Db	799	CAGACGAGCTGGAGGAACCTGTAAGAGCTGTCGCGCAGCGC-----ACCGGATCAG	852
Qy	357	ValGlyIleLeuSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLys	376
Db	853	GTG-----GTCAGCTTCTGAGCGCCATGTCACCGCGCGAGCTTCGCGCGCCGTCAC	906
Qy	377	AsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGly	396
Db	907	GGCTACAGCATCAACCCCAACACAGAAATCGTCGCCCTCGCGCTGGCCACATCGGCGCC	966
Qy	397	SerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsn	416
Db	967	GGGGTCTCCAGGGCTTCGCCCATCGCGCGCGCGCACTCAGCACCGCGGTGAACGACATG	1026
Qy	417	AlaGlyCysLysTyrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeu	436
Db	1027	GTGCGCGCAAGACCCAACTGGTGTGCGGTGTCGCCGCTGGTGTGTCGCGCCACCTG	1086
Qy	437	LeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleVal	456
Db	1087	TTGCTGTGAACAGCCCTGGGTGCGGTCGCGTCCGATGCGCGCTGCGTGGTCTCTGTG	1146

Qy	457	SerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPhelysValAspLys	476
Db	1147	CTGGCCCGCTGGGGCCTGTATCGACGTGCAGCGCCCTTGAAGGCGCTTCTTGAAGCTCAGCCGC	1206
Qy	477	PhaAspPheValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIle	496
Db	1207	TTTGAGTTTCAGCTTCGCTTCGTGACCACTCGGCTGCTCTACGCTCGCGTGTCTGCCG	1266
Qy	497	GlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgPro	516
Db	1267	GGAATCTTCGTCGCGTCAGCATCGCGTACTCGCGCTGTCTTACTACACTATCGCCCG	1326
Qy	517	ArgThrPheValLeuGly-----AsnIleProAsnSerValIleTyrArgAsnVal	533
Db	1327	AGCGACGCGTGTCTGCGCTGGATCGACGCGCATCGACGCGCAGGTCT-----GAGCTG	1377
Qy	534	GluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaPro	553
Db	1378	GCCNAGTACCGCAGGCCACCACCTCTTCGCGGCTGGTGATCTACCGCTTCGACGCGCCC	1437
Qy	554	IleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGlu	573
Db	1438	CTGCTGTTCTTCAACGCCGAGTACTTCAAGCAGCGCGTGTCTTCGCTGGTGGAC-----	1491
Qy	574	GluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSer	593
Db	1492	-----GGCAGCGAGCGCCGAAATCGCGTGTCTTCAACGCGCGAA	1530
Qy	594	AlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLysIleThr	613
Db	1531	GCCATGACCACTCGACATCAGCGGCTGGCCACCCTCCACGAGGTGCACACAGATCCTC	1590
Qy	614	GluArgArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLysLysLeu	633
Db	1591	AAGGCCCAGCGCGTCACCTGTCTGCTGGCGCGGTGACCGGCGACGCTGGACCTGCTG	1650
Qy	634	AsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGlu	653
Db	1651	CAACGCTCGAGCATGCTGGCGAGATCAACGCCCGCTGTGTTCAGCTCGGTGCGCTCC	1710
Qy	654	AlaValGlyAlaCysAsnPheAsnLeuArg	663
Db	1711	GGGTCAGTGCTTATCGCTACTCGCTCGCGC	1740
RESULT 9			
US-09-252-991A-7184			
; Sequence 7184, Application US/09252991A			
; Patent NO. 6551795			
; GENERAL INFORMATION:			
; APPLICANT: Marc J. Rubenfield et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA			
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.136			
; CURRENT APPLICATION NUMBER: US/09/252,991A			
; CURRENT FILING DATE: 1999-02-18			
; PRIOR APPLICATION NUMBER: US 60/074,788			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 60/094,190			
; PRIOR FILING DATE: 1998-07-27			
; NUMBER OF SEQ ID NOS: 33142			
; SEQ ID NO 7184			
; LENGTH: 1767			
; TYPE: DNA			
; ORGANISM: Pseudomonas aeruginosa			
US-09-252-991A-7184			
Alignment Scores:			
Pred. No.:		2,91e-73	Length: 1767
Score:		676.50	Matches: 172
Percent Similarity:		50.17%	Conservative: 117
Best Local Similarity:		29.86%	Mismatches: 242
Query Match:		19.47%	Indels: 45

Alignment Scores:	
Pred. No.:	2,918-73
Score:	676.50
Percent Similarity:	50.17%
Best Local Similarity:	29.86%
Query Match:	19.4%
Length:	1767
Matches:	172
Conservative:	117
Mismatches:	242
Indels:	45

```
DB: 4 Gaps: 13
US-10-762-049-18 (1-680) x US-09-252-991A-7184 (1-1767)
QY 99 TyrThrPheGlnPheLeuLysAlaAspLeuLleAlaGlyIleThrIleAlaSerLeuAla 118
DB 11 TACCGCGCGCCCTGGTTCGCGATGTCACGCGCGGCTGTCGGTAGCGCAATCCAG 70
QY 119 IleProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProPheLeuLeuGlyLeuTyr 138
DB 71 ATCCCAACCGCATCGTACGCCAGATCGCGGCTTCCGCGCCAGGTAGGCGCTCTAT 130
QY 139 SerSerPheIleProPheLeuTyrAlaMetMetGlySerSerArgAspLeuAlaVal 158
DB 131 CGGTGCATCTCGCGATGCTGATCTACGCCCTGATCGCAGCTCGCGCAGTTGATGTA 190
QY 159 GlyThrValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspPro 178
DB 191 GGCCCGACGCGCGC-----ACGCGCGCATGTGCGCGCCGATCACCCG 238
QY 179 -----AsnGluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAla 196
DB 239 CTGCGCGCGCGCATCGCAGCCCTGGTGGACCTGTGATGATCTCGCATCTCGCGATCATGGTC 298
QY 197 GlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSer 216
DB 299 GGGCTGTTCATCTGTCGCGCGCTGGCGCGCGGGTTTCATCGCCAGCTTCCTCTCG 358
QY 217 HisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeu 236
DB 359 CGACCGATCTGTGTCGGTCACTCAACGCGCATCGGCTGAGCTGTGTCGGGCAACTG 418
QY 237 LysSerIleLeuGlyLeuGlnHisPheThrHisGly-----AlaAspIleIleSerVal 254
DB 419 GGCACAGCTGTTCGGCTACGAGCGCGCGACGAGCGGCTTCGTCCCGGCACTCTCGCGCTG 478
QY 255 MetArgSerValPheThrGlnThrHisGluTyrArgTyrPheLeuSerAlaValLeuGlyCys 274
DB 479 CTGGAAACCTGTGT-----CACATCCATTGGCGACGCTGATCTCGGCAGC 526
QY 275 ValPheIlePhePheLeuLeuSerThrArgTyrPheSerIleLysArgProArgPhePhe 294
DB 527 CTCCTCCTCTGTGATGTGTGTGTG-----CCGCGCGCGCTTC 565
QY 295 TrpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuValValTyrPhe 314
DB 566 ----CCGAGTGGCCCGGCGCTCTCGGGGCTCTCGGCCAGCTCTCGCGCGGCGTG 622
QY 315 ThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnPro 334
DB 623 CTGGGCTTGGATCGCTACGCGTGGAACTGCTTGGCGAGGTACCGCGCGGCGTG----- 676
QY 335 ProSerLeu-----ThrAsnLeuValPheValSerProTyrMetThrThrAla 350
DB 677 CGCAACTAGCTAGCTGGCGCAGACCACTGGAGGAACTGAAAAAGCCCTGTCGCGCAGCGC 736
QY 351 ValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArg 370
DB 737 -----ACCGCATACGGTG-----GTGAGCTTCTGAGCGCCCATGCTACCGCGCGC 784
QY 371 SerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGly 390
DB 785 AGCTTCGCGCGCGCTCACGCTACAGCATCAACCCCAACACAGAAATTCGTGCCCTCGGC 844
QY 391 ThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArg 410
DB 845 CTGGCCAAATCATCGCGCGCGGTCTCCAGGGCTTCGCCATCAGCGCGCGCCGACCTACGC 904
QY 411 SerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeu 430
DB 905 ACCGCGGTGAACGACATGTGCGCGCGCAGACCCAACTGTCGGCGTGTGTCGCGCGCTG 964
QY 431 AlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValVal 450
```

```
DB 965 GTGATCGCGCGCACCCCTGTGTCTGTGAACAAGCCCTGGGTGGGTGCGGATGCGCGG 1024
QY 451 LeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaIleHis 470
DB 1025 CTCGTGCGGCTCTGTGCTGGCGCGCTGGGCGCTGATCGACGTGCGCGCTGAAGGCG 1084
QY 471 LeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValVal 490
DB 1085 TTCTGGAAGCTCAGCGGCTTCAGTTTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTC 1144
QY 491 PheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeu 510
DB 1145 AGCGTCGCGCTGCTCGCGGAATCTTCGTCGCGCTGAGCATCGCGTACTGCGCTGCTC 1204
QY 511 LeuPheIleAlaArgProArgThrPheValLeuGly-----AsnIleProAsnSer 527
DB 1205 TACTACACCTATCGGCCGAGCAGCGGTGCTGCGCTGCGATCGACGCGCATCGCGCCAG 1264
QY 528 ValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIle 547
DB 1265 GTC-----GAGCTGGCCAAGTACCGCAGGCCACCACTCTTCCCGCGCTGTGATC 1315
QY 548 LeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThr 567
DB 1316 TACCGCTTCGACCGCCCTGCTGTCTTCAACGCGGAGTACTTCAAGCAGCGGTGCTT 1375
QY 568 ArgTrpIleAspGluGluGluArgIleLysAlaThrGlyLysThrSerLeuGlnTyr 587
DB 1376 GCCGTGGTGAGC-----GGCAGCGAGCGCGCGAATGCG 1408
QY 588 ValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607
DB 1409 GTGCTGCTCAACGCGCAGCCATGACCACTCGACATCAGCGGCTGGCCACCTCCAC 1468
QY 608 GluValLysLysIleThrGluArgGluLeuGlnLeuValLeuValAsnProValSer 627
DB 1469 GAGGTGCAACAGATCTCAAGGCCCGCGGTGCACTGTCTGCTGGCGCGGTGACCGGG 1528
QY 628 GluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysIleTrpIle 647
DB 1529 CAGACGCTGGACCTGCTGCAAGCTCGAGCATGTCTGGCGAGATCAAGCGCGCTGGTG 1588
QY 648 TyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArg 663
DB 1589 TTCAGCTCGGTGCGTCCGGGTCAGTGCCTATCGTACTGCTGCTGCGC 1636

RESULT 10
US-09-720-317A-5
; Sequence 5, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-5

Alignment Scores: 2,05e-73 Length: 1240
Pred. No.: 675.50 Matches: 132
Score: 67.22% Conservative: 69
```



```
Best Local Similarity: 44.15% Mismatches: 97
Query Match: 19.44% Indels: 1
DB: 4 Gaps: 1

US-10-762-049-18 (1-680) x US-09-785-381-5 (1-1240)

QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
DB 6 AGCTCTACGGAGGATCGCGTGGCGATCTTCGCCCTCCGTAAGAGGGTACAGACTC 65
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
DB 66 GACGCAACAGAGAGTCTGCCCATGGGTTCCTCCCAACGTTGCTGGTCTCTCTCTCTCG 125
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
DB 126 TGCTATGTGCACACAGGTTCTTCTCCGCAACGCGAGTGAACCTTCAGCGGGGGCCAGG 185
QY 421 ThrAlaAlaSerAsnIleMetSerLeuAlaValMetLeuThrLeuPheLeuThr 440
DB 186 TCGACCGTTTCAACATCGTCTCATCGTCCATCCCGTGTTCGTCACCTGGAGCTGTTCATG 245
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
DB 246 AAGCTCTCTACTACACGCCCATCGCGTGTCTGCCCTCCATCATCTCTCGGCTCTTCG 305
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
DB 306 GGACTGATCGACATCAAGAGGCGCTGCAGCATATGGAAGATCGAAGATGGATTTCCCTC 365
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
DB 366 ACCTGCTCGGTGTGTGTGGTCTCTGTGGTGGTGGAGATGGGCTTGCAGTT 425
QY 501 AlaIleValIleSerValLeuAlaValLeuPheIleAlaArgProArgThrPheVal 520
DB 426 GCATTGGCATTTCTTCGCAAGATCATCATACAGTCGCTTCGGCTCAGTGGAGATC 485
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
DB 486 CTTGGCAGGCTTACAAGGACAGATATCTTCTGCAGCGTCAGGAGTACCCCTGTAGCTGC 545
QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaPro---IleTyrPheAlaAsnAla 559
DB 546 CTAACCTCGACTGTACTGCTTATACGGTGCACATCTTCTCTCTCTCATCAAGCC 605
QY 560 SerTyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAla 579
DB 606 ACTTCCGCTCAAGAAAGGATCACAGAGTGGTGGTGGAGAGGAGTGGAGCTCAATGGA 665
QY 580 ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp 599
DB 666 AAAGCGAGGAGAGATACAGAGGATGTCTTGTATGATGCAAGTGTGGTAAACATCGAC 725
QY 600 ThrSerGlyIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGln 619
DB 726 ACTTCAGGACTCACTGCACTGGAAGAAATACAAAGAGTGGTGTCTCTTGGCTTACAG 785
QY 620 LeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln 639
DB 786 ATGCTATAGCCATCGCGGTGGAAGGAGTTCAGAGATGAAGATGTCACAGGTGGTG 845
QY 640 AsnHisLeuGlyLysIleTyrIleThrValGluGluAlaValGlyAlaCys 658
DB 846 GACAGGGTAGGACAGGATGGATCTTCATGACAGTAGGTGAAGCGTGGAGGCGCTGT 902

RESULT 11
US-09-785-381-4
; Sequence 4, Application US/09785381
; Patent No. 6602992
; GENERAL INFORMATION:
; APPLICANT: DALLOS, Peter
; APPLICANT: ZHENG, Jing
```

```
APPLICANT: MADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-3701
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 2441
; TYPE: DNA
; ORGANISM: Mus sp.
; US-09-785-381-4

Alignment Scores:
Pred. No.: 7,8e-70 Length: 2441
Score: 651.00 Matches: 191
Percent Similarity: 45.95% Conservative: 144
Best Local Similarity: 26.20% Mismatches: 270
Query Match: 18.73% Indels: 124
DB: 21

US-10-762-049-18 (1-680) x US-09-785-381-4 (1-2441)

QY 35 MetAsnAsnPheGluArgValHisGlnValGluValProPro----- 48
DB 198 CTCGACAGCATGGATCATGCTGAAGAAATCCCTGCAGACGCCAGAGGTACTAC 257
QY 49 ---ProGlnProPhePhe-----LysSerLeuLysTyrSerLeuLys 61
DB 258 GTGGAAGGCCCATCTTTCAGTCATCTCTCTCAAGAGAGGCTGCACGTCAGGACAAA 317
QY 62 GluThrPhePheProAspAspProLeuArgGln---PheLysAsnLysProAlaSerLys 80
DB 318 GTCACAGAGTCCATTGGAGATAAGCTGAAGCAGGATTCACGTGTACTCCT---AAA 371
QY 81 LysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThr 100
DB 372 AAAATAAGAAACATCATTTACATGTTCTCTGCTTATCACTAAGTGGCTGCCAGCATATAA 431
QY 101 Phe---GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIle 119
DB 432 TTCAGAGGATGTGTGTAGTGACTTGTCTCGGATATAGCACTGGGGTACTCCAGCTT 491
QY 120 ProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSer 139
DB 492 CCCCAAGGCTTAGCCTTCGCCATGCTGGCAGCGCTGCTCCGGTGTGGCTGTACTCA 551
QY 140 SerPheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGly 159
DB 552 TCGTTTTTACCCCGTTATCATGTACTGTTCTTTTGAACCTTCAAGACACATATCTATAGT 611
QY 160 ThrValAlaValGlySerLeuMetGlySerMetLeu----- 172
DB 612 CCTTTTGTCTTATTAGTTGATGTGGAGTGGTGGCCCTCGCTGTAGTACCAGATGAT 671
QY 173 -----SerAsnAlaValAspProAsnGlu---AspProLysLeuTyr 185
DB 672 ATTGTATCCAGGAGGAGTAAATGCAACCAACGGGACAGAGCCAGAGATGCCTAAGA 731
QY 186 LeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeu 205
DB 732 GTGAAAGTCGCCATGTCTGTACTTCTTACAGGAATCATTCAGTTTCCCTAGGTGTC 791
QY 206 PheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGly 225
DB 792 TGTTAGTTGGATTTGTGGCCATATACCTCAGGAGCCATTTGGTGGAGCTTTTACCAC 851
QY 226 GlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPhe 245
DB 852 GCGGCTGCTGCCACGTTTTCACGTCCATGTTTAAATACCTGTTTGGGGTCAAAACAAAG 911
```

```

Db      1950 GTTGGAAATGCCAACGTGGCCCAATGCTACTGTGTCTCAAGTGGATGCAGAAAGTAGACGGA 2009
Qy      572 -----GluguGluGluarg-----Ile 577
Db      2010 GAAATGCTACAAAACCTGAAGAAGAGGATGATGAAGTCAAATTTCCCCCAATAGTCACT 2069
Qy      578 LysAlaThrGlyGluThrSerLeuGln-----Tyr 587
Db      2070 AAAACACACATTTCTCTGAAAGAGCTGCGAGAGATTTTGTCCCAGGGGAAAAATGTCCACACT 2129
Qy      588 ValIleileAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607
Db      2130 GTCACTTAGACATTCTAGCAGAGTCAATTTTGTGGATTCGTCTGGAGTGAACCTGTGCC 2189
Qy      608 GluValbyslystleThrGluArgGluLeuGlnLeuValLeuValAsnProValSer 627
Db      2190 GCGATTGTGAAAGAATATGAGATGTTTGGAAATTTATGTATATTATGACGAGATGCAGCCCA 2249
Qy      628 GluValMetLysLysLeuAsnLysSerLys--PheGlnAsnHisLeuGlyLysLysTrp 646
Db      2250 CAAGTTGTGAATGACCTCACCCGCAACACTTTTTTGAATAATCCTGCCTTTGAAAGAGCTT 2309
Qy      647 IleTyrLeuThrValGluGluAlaVal 655
Db      2310 CTGTTCCACAGTATCCAGATGCAGTC 2336

RESULT 12
US-09-785-381-2
; Sequence 2, Application US/09785381
; Patent No. 6602992
; GENERAL INFORMATION:
; APPLICANT: DALLOS, Peter
; APPLICANT: ZHENG, Jing
; APPLICANT: MADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-37UL
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 4113
; TYPE: DNA
; ORGANISM: Meriones unguiculatus
US-09-785-381-2

Alignment Scores:
Pred. No.: 2,66e-69 Length: 4113
Score: 650.00 Matches: 189
Percent Similarity: 46.63% Conservative: 143
Best Local Similarity: 26.54% Mismatches: 254
Query Match: 18.71% Indels: 126
DB: Gaps: 21

US-10-762-049-18 (1-680) x US-09-785-381-2 (1-4113)

Qy      39 GluArgValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyr 58
Db      311 GAGAGGCTGCAT-----GTCAGAGAC 331
Qy      59 SerLeuLysGluThrPhePheProAspAspProLeuArgGln---PheLysAsnLysPro 77
Db      332 AAAGTCTCAGAGTCCATT-----GGGGATAAAGCTGAAGCAGCGGTTTCATGCACCTCCC 385
Qy      78 AlaSerLysLysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTTPAlaPro 97
Db      386 -----AAAAGAATAGAAACATCATTTACATGTTCTGCCCCATCACTAAGTGGTGGCCA 439
Qy      98 LysTyrThrPhe---GlnPheLeuLysAlaAspLeuIleAaGlyIleThrIleAlaSer 116

```

Db 440 GCTTACAGTTCAAGGAGTATGTGTGGTGACTTGGTTTCAGGCATAAGCACCGCGTG 499
Qy 117 LeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAlaSerLeuProIleLeuGly 136
Db 500 CTTGAGTCTCCCAAGGCTTACGCTTGCATGCTCGCGCTGTCTCTCGGTTTCGCG 559
Qy 137 LeuTyrSerSerPheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeu 156
Db 560 CTGACTCTTCATTTTATCTCTGTTATCATGACTGTTCTTCTTGGGACCTCCAGACATA 619
Qy 157 AlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMetLeu----- 172
Db 620 TCTATAGTCTCTTCGCGCTCATTTAGTTTGTATGATCGTGTGTGGTGCTCGCGTGTGTC 679
Qy 173 -----SerAsnAlaValAspProAsnGlu-----AspPro 182
Db 680 CCGCATGACATCGTCATCCCGGAGAGTGTAACCAACCAACGCGACGCGCCGAGAC 739
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 740 GCGCTGAGTGAAGTCGCTGTCACCTGCTCTCAGGAATCATTCAGTTTTCG 799
Qy 203 LeuGlyLeuPheAlaLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGly 222
Db 800 CTAGGTGTGCAAGTTGGATTGTGGCCATATACCTCAGGAGCGCTGGTCGAGGG 859
Qy 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeu 242
Db 860 TTACACCGCGCGCGCTGCTCTTACATCCATGTTGAAATACCTGTTTGGGGTT 919
Qy 243 GluHisPheThrHisGlyAlaAspIleIleSerValMetArgSer-----ValPhe 259
Db 920 AAGCAAAAGCGGTAC---AGTGGATCTTTTCGGTGTATATATACAGTTGCTGTGTG 976
Qy 260 ThrGlnThrHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePhe 279
Db 977 CAGATGTTAAAACTCAAGCTGTCTCCCTAGGCGTGGCGTGAATGTTTGGTTG 1036
Qy 280 LeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheTyrValSerAlaMet 299
Db 1037 CTGTTGGTGGCAAGGAGTTTAATGAGAGA-----TTTAAAGAAATTCGCCA 1084
Qy 300 AlaProLeuThr-----SerValIleLeuGlySer---LeuLeuValTyrPhe 314
Db 1085 GCACCCATCTCTPAGAGTTCTTTGCTGTGTGTGGTATGCGAATCGCATTTCCGCGGGGTTT 1144
Qy 315 ThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnPro 334
Db 1145 AACTTGACAGTCTCTACAGTGTGGATGTGTGGAACTCTTCTCTGGGGCTACTCCCT 1204
Qy 335 Pro-----SerLeuThrAsnLeuValPheValSerProTyrMetThrThr 349
Db 1205 CCTGCCAACCGGACACCGCTCTTCCACTCGTGTATGTGGAT----- 1249
Qy 350 AlaValLysThrGlyIleValValGlyIleSerLeuAlaGluGlyIleAlaValGly 369
Db 1250 -----GCCATTCGATGACCATCGTGTGGATTTTTCAGTGACAATTTCCATGGCC 1297
Qy 370 ArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIle 389
Db 1298 AAAACCTTGGGGAATAAGCATGCTCTCTTCCAGACCTTCTCCATTTCTCTGCTCTGTCT 1357
Qy 390 GlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSer 409
Db 1358 GGGATATGCAACTCCATCGGATCTCTCTTCCAGACCTTCTCCATTTCTCTGCTCTGTCT 1417
Qy 410 ArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleLeuMetSer 429
Db 1418 CGCAGCCTTGTTCAGGAGGAACTGGAGGAAAAACACAGCTCGCAGGTGCTTGGCGCTCG 1477
Qy 430 LeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuVal 449
Db 1478 CTGATGATCTGCTGTCATTTTATGACCATCTGGATTCCTCTTGTAGTCATTTGCCCCAGGCT 1537

Qy 450 ValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu---IleAspTyrGluAlaAla 468
Db 1538 GTGCTCTCGGCCATTTGATCGTCAACCTGAAAGGATGTTTATGCAAGTCTCAGACTGTG 1597
Qy 469 IleHisLeuPheLysValaAspLysPheAspPheValValCysMetSerAlaTyrIleGly 488
Db 1598 CCCTTCTTCTGGAGAACCCAGCAAAATAGACTGACCATCTGGCTTACACCTTTGTGTCC 1657
Qy 489 ValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArg 508
Db 1658 TCCCTGTCTCTGGCTTGGACTACGGACTGATTACTGTCTGTGATCATTTGCTGTGACT 1717
Qy 509 ValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerVal 528
Db 1718 GTGATTTACAGAACCCAGAGTCCGAGCTCAAGGTCTCTGGGCGAGCTCCCTGCACCGAT 1777
Qy 529 IleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeu 548
Db 1778 GTATACATTGACATAGACGCATATGAGGAGTGAAGAAATTCCTGGAATAAAATATTC 1837
Qy 549 GluIleAspAlaProIleTyrPheAlaAsnAla-----SerTyr 561
Db 1838 CAGATAAACGCCCAATTTACTATGCAACAGTGACTTGTGTATAGCAACGCCCTAAAAAGA 1897
Qy 560 ----- 1897
Db 1898 AAGACTGTGTGAACCCAGCGCTCATATGGGAGCAAGGAAGGCCATGAGGAAGTAC 1957
Qy 562 LeuArgGlu-----ArgIleThrArgTyr 569
Db 1958 GCAAGGAAGTTCGGAAACGCCAACATTTGCCAACGCAGCTGTTGTCAAAGTGGATGGAGAA 2017
Qy 570 IleAsp-----GluGluGluGluArg----- 576
Db 2018 GTAGATGGAGAAATCTCTAGAACGCCGAAGAGAGAGATGATGAAGTAAATATATCCCCA 2077
Qy 577 -----IleLysAlaThrGlyGluThrSerLeuGln----- 586
Db 2078 ATAGTCATCAAAACAATTTCTCTGAAGAGCTCGACAGATTTTATGCCCCAGACAGAAAT 2137
Qy 587 -----TyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSer 604
Db 2138 GTCACACATCATCTAGACTTCACACAGTCAATTTTATCGACTCTGTGTGAGTAAAA 2197
Qy 605 MetLeuGluGluValLysLysIleThrGluArgGluLeuGlnLeuValLeuValAsn 624
Db 2198 ACCTGCTGTGTGTGTGAGGATACGGAGATGTTGGTATTTATGTGTACTTAGCAGGA 2257
Qy 625 ProValSerGluValMetLysLysLeuAsnLysSerLys---PheGlnAsnHisLeuGly 643
Db 2258 TGCAGCCCAACAGTCTGTAATGACCTCACCCGCAACCGTTCTTTGAAATCTCGCTTA 2317
Qy 644 LysLysThrIleTyrLeuThrValGluGluAlaVal 655
Db 2318 AAAGAGCTCTGTTCACAGATATCCATGATGCAGTC 2353

RESULT 13

US-09-902-540-7746
; Sequence 7746, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825

[illegible]

Db	862	GGCGGATTTCACGCCGACAGGTTCCGCTACCGGCTGCACACCAACGAGGATTCCTCGGG	921
Qy	389	IleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPhe	408
Db	922	CAGCGCCGACGCCAACCTCGCCACTCGCGTTTCAGCCAGGGGTTCCCGCTGACGGGAGCGAC	981
Qy	409	SerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMet	428
Db	982	TCGCGCACGCCCTCAACGCTCCATATGGGGGGCGGACCCAGCTCGTGGGGGTCTCGCC	1041
Qy	429	SerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeu	448
Db	1042	GCAGGGGTGTGTGGTGTTCGCTTCTCCTCAGCCCTGTCTCCACGACTTGCCTCATG	1101
Qy	449	ValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAla	468
Db	1102	GTACGCTGGCGCCATCGCTTCGTGGCCCGCGGTGACCTGTGTGGAGTTCCGAGCATC	1161
Qy	469	IleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGly	488
Db	1162	ATCGACTGTGGCGGTGGCGGGTGGAGCGGTGCTCGCGTCCGTGACATGGCGGGG	1221
Qy	489	ValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArg	508
Db	1222	GTGCTGTGTGGCATCTCTCAAGGCATCTCTGTCGCGTGGCGCTTGCCTTGCAGC	1281
Qy	509	ValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerVal	528
Db	1282	CTCATCGCGCGGCCCGCCAGGCCCCACGACGCGGTGCTTGGCGAGCGCGGTGCC	1341
Qy	529	IleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeu	548
Db	1342	GGCTACACAGCATCGACGGGTTTCGAGACGCCCGAGACGCTGTCGCGGCTCGTCATCTAC	1401
Qy	549	GluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArg	568
Db	1402	CGCTTCGATCGCGCGCTGTTCTTCGCAACGCGCGCATCTGCGG	1446
Qy	569	TrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrVal	588
Db	1447	-----GAGCAGGCCCGGGCTTGGTCTCCAGCGCGATCGCCGCTGCGGTGTTTC	1497
Qy	589	IleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGlu	608
Db	1498	GTGCTTGATCGTCCCGCGTTTCGACATGGACGTACCCCTCCGATGCCCTGGAGAG	1557
Qy	609	ValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSerGlu	628
Db	1558	CTCGCCCGTGAAGTTCAGGAAGAGGCGTGGTCTGCGCGTCCCGAGGCTCGGCGCGC	1617
Qy	629	ValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIleTyr	648
Db	1618	CTGCGCGCGTGTGGCGCGTACCGGCTCTGTGAGCGGCTGTGGAGCGGTGGAGATGTGCAT	1677
Qy	649	LeuThrValGluGluAlaVal	655
Db	1678	GCCACCGTGGAGCGCGGTG	1698
RESULT 14			
US-09-902-540-768			
; Sequence 768, Application US/09902540			
; Patent No. 6833447			
; GENERAL INFORMATION:			
; APPLICANT: Goldman, Barry S.			
; APPLICANT: Hinkle, Gregory J.			
; APPLICANT: Slater, Steven C.			
; APPLICANT: Wiegand, Roger C.			
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof			
; FILE REFERENCE: 38-10(15849)B			
; CURRENT APPLICATION NUMBER: US/09/902,540			
; PRIOR FILING DATE: 2001-07-10			
; PRIOR APPLICATION NUMBER: 60/217,883			

; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 768
; LENGTH: 4854
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-768

Alignment Scores:

Pred. No.: 2,316-68 Length: 4854
Score: 643.50 Matches: 162
Percent Similarity: 50.44% Conservative: 124
Best Local Similarity: 28.57% Mismatches: 256
Query Match: 18.52% Indels: 25
DB: 4 Gaps: 7

US-10-762-049-18 (1-680) x US-09-902-540-768 (1-4854)

QY 91 ProIlePheGluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAspLeuIleAla 110
DB 68 CCGGGCGCTGCGGAGGACCGCGCTATCAGCGCGCTGGTTCCGGGGGAGACCTGCTCG 127
QY 111 GlyIleThrIleAlaSerLeuAlaIleProGlnGlyLeuSerTyrAlaLysLeuAlaAsn 130
DB 128 GCCTACCAATTGGTGGATGCTATCCCCAGGGCTGACCTACGGCGAGCTCGTGGC 187
QY 131 LeuProPheLeuGlyLeuTyrSerSerPheIleProProLeuLeuTyrAlaMetMet 150
DB 188 GTGGTCCAGAGCGGGCTCTACGGGGCGTGGGAATGCTGGCCTATGCGCTGTTTC 247
QY 151 GlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySer 170
DB 248 GGGCCCTCGCGGACCTCATCGGACCCGAGCGGGTCCGCCATCTCCACCGCGCG 307
QY 171 MetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAlaPhe 190
DB 308 GCCTTCGCG--CGGTGGCGGGGGGGCGGCCCGCGCGCTACCGATCGCTGGCGGCA 364
QY 191 ThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuGlyLeuPheArgLeuGlyLeu 210
DB 365 CTGCTGCGCTGCTGGTCGCGGTGTGAGCGTCTCGGAGGGTGTCTCAAGGTGGCGG 424
QY 211 IleValAspPheLeuSerHisAlaThrIleGlyPheMetGlyGlyAlaAlaThrVal 230
DB 425 CTCGCGGACTTCTGTCTCAAGCCATCTCTATCGGCTACATCAATGCGCGGCGCTCATC 484
QY 231 ValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis-----PheThrHisGly 248
DB 485 ATCATGGGAGCGAGCTCGCGCGCTCTTCGGGAGGAGCGCCAGTCCGACACGTTCTCG 544
QY 249 AlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGlu 268
DB 545 GGTCAAGTGTTCGAGGTGGCCACCCACCTC--GAAGGAGCCACCGTC-----CCG 592
QY 269 SerAlaValLeuGlyCysValPheIlePhePheLeuSerThrArgTyrPheSerLys 288
DB 593 ACCTCTCTGCTGGGCTGGCGGCTCATCACCGCGCTGCTGTGTGGGCAACTCTCGGCC 652
QY 289 LysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGly 308
DB 653 AAGTGCCC-----GGTCCGCTCATCTCGTGGTGGTCCCTACC 688
QY 309 SerLeuLeuValTyrPheThrHisAlaGlyLeuHisGlyValGluValIleGlyGluLeu 328
DB 689 ACGTGGCGGAGGGCTGTTCCAGTTGAGCATGGGGCATCAAGTGTGGGGGCCCTC 748
QY 329 LysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThr 348
DB 749 GCGCGCGAGCCCCCGCTCCAGGCGCTCGCTCCCTGCGTCTCGAGGAC-----GTGCGG 802
QY 349 ThrAlaValLysThrGlyIleValValGlyIleLeuSerLeuAlaGluGlyIleAlaVal 368
DB 803 TCCTGCTCTCCCGCGCTTCAGCCTTGGCGCTCGTCAACTACGCCAGCTCCGTTGTCAGC 862

QY 369 GlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAla 388
DB 863 GCGCGATTCTACGCCGACAGGTTCCGCTACCGCTGGACGACGACCAACGAGGATTCCTCGG 922
QY 389 IleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPhe 408
DB 923 CAGGCCGAGCCCAACCTCCGCTGTCAGCCAGGGGTTCCTCCGTGAGGAGCGAC 982
QY 409 SerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleMet 428
DB 983 TCGCGCAGCGCGCTCAACGCTCCATGGGGGGGCGGACCCAGCTCGTGGGGGTGCTCGCC 1042
QY 429 SerLeuAlaValMetLeuThrLeuPheLeuThrProLeuPheHisTyrThrProLeu 448
DB 1043 GCAGGGGTGTGTGTGCTTGTCTCCTCACGCCCTGCTCACGACTTGTGCCCATG 1102
QY 449 ValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAla 468
DB 1103 GTGACCTGGGGCCCATGCTTCTGGCGCGCGGTACTCTGTGAGTTCGAGGCTC 1162
QY 469 IleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGly 488
DB 1163 ATCGACCTGTGGCGCTGCGCGCGTGGAGCGGTCTCGCGTGGTACGATGGCGGG 1222
QY 489 ValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArg 508
DB 1223 GTGCTGTGTGGGATCTCTCAAGGATCTCTCGTGGGTGGCGCTTGGCGAC 1282
QY 509 ValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerVal 528
DB 1283 CTATCCCGGGCGCCAGGCCCCAGAGCGGTCTTGGCGAGCGCGGCGGTGCCCC 1342
QY 529 IleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeu 548
DB 1343 GGCTACACGACATCGAGCGGTTCGAGACGCGGACGCGGTGCGCGCTCGTCATCTAC 1402
QY 549 GluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluIleThrArg 568
DB 1403 CGCTTCGATGCGCGCTGTTCTTCGCAACGCGCGCATCTGCGG-----1447
QY 569 TrpIleAspGluIleGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrVal 588
DB 1448 -----GAGCAGCCCGCGGCTTGGTCTCCAGCGGATGCGCCCGTGGTTC 1498
QY 589 IleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGlu 608
DB 1499 GTGCTTGATGCGTCCGCGTGTTCGACATGGAGTCAACCGCTCCGCTCGGAGAG 1558
QY 609 ValLysLysIleThrGluArgGluLeuGlnLeuValLeuValAsnProValSerGlu 628
DB 1559 CTGCGCGTGTGAGTTCGAGGAAGAGGCGTGTGTGGGTGTCGCGGCTCGCGGCGC 1618
QY 629 ValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTyrIleTyr 648
DB 1619 CTGCGCGCGCTGCTGGCGGCTCCCGCTGTCGAGCGGTGGCGCGCGGAGATGTGCAT 1678
QY 649 LeuThrValGluGluAlaVal 655
DB 1679 GCCACCGTGGAGCGCGGTG 1699

RESULT 15

US-09-949-016-4516
; Sequence 4516, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4516
; LENGTH: 2832
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4516

Alignment Scores:
Pred. No.:      1,55e-68      Length:      2832
Score:          641.50       Matches:     184
Percent Similarity: 44.92%    Conservative: 143
Best Local Similarity: 25.27% Mismatches:     290
Query Match:      18.46%     Indels:        11
DB:               4         Gaps:         22

US-10-762-049-18 (1-680) x US-09-949-016-4516 (1-2832)

Qy   18 AsnThrIleIlcGluAspMetGlySerValAspTyrGluTyrProLeuGlyMetAsn--- 36
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   55 AAGCTTTCAACCCAGAGACTCAGCTGAAGAAATGACAGTTATTCATCGGTGGGATCCATCTG 114

Qy   37 -----AenPheGluArgValHisGlnValGlu 45
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   115 GAACCTCAAAGGAATCAAGTACTGACTTCAAGCAATTGGACCAATGATCAATGC--- 171

Qy   46 ValProProGlnProPhePheLysSerLeu-----LysTyrSerLeuLysGluThr 63
    :|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db   172 -----AGACCITATCATGATCCTTATTGAGCGGTCAAGAGAATCAGATACA 219

Qy   64 PhePheProAspProLeuArgGlnPhe----LysAsn-----LysProAla--- 78
    ||| ::| :|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db   220 AACTTCAGGAGTTTGTTATTAAAAAGCTGCAGAGAATAATGCCAGTCAGTCAGCCAAA 279

Qy   79 SerLyLePheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLys 98
    |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db   280 GCCAAAATATGATTTAGT-----TTCCTCTCTGTTTGCAGTGGCTCCCAAAA 330

Qy   99 TyrThrPhe--GlnPheLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeu 117
    |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db   331 TAGGACCTAAAGAAAACATTTTAGGGATGTGATGCAGGCTTGATGTGGGCATATTTA 390

Qy   118 AlaIleProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeu 137
    :|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db   391 TTGTGGCCCCAGTCCATTGCTTATTCCTGTGCTGGCCCAAGAACCTGTCTATGGTCTG 450

Qy   138 TyrSerSerPheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAla 157
    :|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db   451 TACACATCTTTTTTGGCCAGCATATTATTTTCTCTGGGTACCTCCCGTCACATCTCT 510

Qy   158 ValGlyThrValAlaValGlySerLeuLeuMetGly----- 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   511 GTGGGCATTTTGGAGTACTGTGCTTATGATTTGGTGAGACAGTTGACCGAGAACTACAG 570

Qy   169 ----- 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   571 AAAAGCTGGCTATGACAATGCCCATAGTCTCTTCTTAGGAATGGTTTCAAAATGGAGC 630

Qy   170 SerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTy---LeuHisLeu 188
    :|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db   631 ACATATTAAATCATCATCAGCAGAGGATATGTGACAAAAGTTGCTATGCAANTATGGTT 690

Qy   189 AlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeu 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   691 GCAGCATCTGAACCTTTATAGCTGGAGTTTATCAGGTAGCGATGGGCTCTTTTCAAGTG 750

Qy   209 GlyLeuIleValAspPheLeuSerHisAlaThrIleLeuGlyPheMetGlyValAlaLa 228
    ||| ::| :|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
    :|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
```

```
QY 569 -----TrpIleAspGluGluGluGluArgIleLysAlaThrGlyGluThr--- 583
Db 1852 ATAAAGTGTGCTTGGAAAGAGGAGCAAGAGAAAGATCAAGAAAAAGTAGTCACTCTT 1911
QY 584 -----SerLeuGlnTyr 587
Db 1912 GGTGGAATCCAGGATGAAATGTGAGTGCACACTTTCCCATGATCCCTTGGAGCTGCATACT 1971
QY 588 ValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607
Db 1972 ATAGTGTACTGCTGAGTGCATTCATTTTAGATACAGCAGGGATCCACACTGAAA 2031
QY 608 GluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSer 627
Db 2032 GAAAGTTGCGCAGAGATTATGAAGCCATTGGAATCCAGGTTCTGCTGGCTCAGTGCATCCC 2091
QY 628 GluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIle 647
Db 2092 ACTGTGAGGGATTCCCTAACCAAGGAGAGATATTGCAAAAGGAAGAAACCTTCTC 2151
QY 648 TyrLeuThrValGluGluAlaVal 655
Db 2152 TTCTATAGTGTGTATGAAGCATG 2175
```

Search completed: August 31, 2005, 16:06:59
Job time : 386 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 31, 2005, 11:53:46 ; Search time 1127 Seconds
(without alignments)
3947.916 Million cell updates/sec

Title: US-10-762-049-18
Perfect score: 3475
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRAKTNPKDTEGNNV 680

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
	Xgapop 10.0 , Xgapext 0.5
	Xgapop 6.0 , Xgapext 7.0
	Delop 6.0 , Delext 7.0

7331713 seqs, 3271544945 residues

Searched:

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10762049/runat_3082005_121246_6030/app_query.fasta_1.839
-DB=Published Applications NA -QFMT=fastp -SUFFIX=rnnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10762049 @CGN 1 1 480 @runat_3082005_121246_6030
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Published Applications NA:**
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
 - 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
 - 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
 - 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
 - 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
 - 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
 - 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
 - 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
 - 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
 - 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
 - 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2595.5	74.7	1977	9	US-09-938-842A-1502	Sequence 1502, A
2	2595.5	74.7	1977	11	US-09-938-842A-1502	Sequence 1502, Ap
3	2330.5	67.1	2372	18	US-10-437-963-38184	Sequence 38184, A
4	2036	58.6	2008	18	US-10-424-599-131116	Sequence 131116, A
5	1856	53.4	2727	20	US-10-425-115-50601	Sequence 50601, A
6	1845.5	53.1	2651	20	US-10-425-115-182207	Sequence 182207, A
7	1844.5	53.1	2366	18	US-10-424-599-65103	Sequence 65103, A
8	1773.5	51.0	2780	19	US-10-437-963-95853	Sequence 95853, A
9	1762	50.7	2656	18	US-10-437-963-77005	Sequence 77005, A
10	1742	50.7	2713	20	US-10-425-114-15719	Sequence 15719, A
11	1734.5	49.9	2429	19	US-10-425-115-75763	Sequence 75763, A
12	1698	48.9	4390	19	US-10-437-963-73410	Sequence 73410, A
13	1676.5	48.2	4390	19	US-10-437-963-77012	Sequence 77012, A
14	1676.5	48.2	1518	18	US-10-424-599-131117	Sequence 131117, A
15	1563.5	45.0	2251	19	US-10-437-963-72062	Sequence 72062, A
16	1513.5	43.6	1980	17	US-10-260-238-1066	Sequence 1066, Ap
17	1500	43.2	2336	18	US-10-424-599-52175	Sequence 52175, A
18	1460	42.0	1992	18	US-10-425-114-13333	Sequence 13333, A
19	1446.5	41.6	3252	19	US-10-437-963-78357	Sequence 78357, A
20	1436	41.3	3201	20	US-10-425-115-44188	Sequence 44188, A
21	1406	40.5	2242	20	US-10-425-115-40196	Sequence 40196, A
22	1379.5	39.7	2506	18	US-10-424-599-74393	Sequence 74393, A
23	1218	35.1	2156	18	US-10-424-599-60281	Sequence 60281, A
24	1215	35.0	2156	20	US-10-739-930-3837	Sequence 3837, Ap
25	1204	34.6	1809	18	US-10-424-599-83137	Sequence 83137, A
26	1185.5	34.1	1632	18	US-10-424-599-74389	Sequence 74389, A
27	1175.5	33.8	2913	19	US-10-437-963-76453	Sequence 76453, A
28	1134.5	32.6	2346	20	US-10-425-115-22539	Sequence 22539, A
29	1016	29.2	1303	18	US-10-424-599-139734	Sequence 139734, A
30	1010.5	29.1	1551	18	US-10-424-599-54350	Sequence 54350, A
31	1009	29.0	1247	18	US-10-424-599-130951	Sequence 130951, A
32	1002	28.8	2058	9	US-09-938-842A-717	Sequence 717, App
33	1002	28.8	2058	11	US-09-938-842A-717	Sequence 717, App
34	955.5	27.5	2988	19	US-10-437-963-92700	Sequence 92700, A
35	934	26.9	680	18	US-10-424-599-32307	Sequence 32307, A
36	923	26.6	1019	19	US-10-767-701-11661	Sequence 11661, A
37	901	25.9	1174	20	US-10-425-115-75765	Sequence 75765, A
38	900	25.9	536	18	US-10-424-599-32308	Sequence 32308, A
39	898	25.8	674	19	US-10-767-701-9435	Sequence 9435, Ap
40	892.5	25.7	1115	19	US-10-767-701-9435	Sequence 9435, Ap
41	887	25.5	1069	20	US-10-425-115-3619	Sequence 3619, Ap
42	814	23.4	1802	19	US-10-437-963-7407	Sequence 7407, Ap
43	737.5	21.2	1018	18	US-10-424-599-18417	Sequence 18417, A
44	736.5	21.2	841	18	US-10-424-599-135930	Sequence 135930, A
45	715	20.6	788	20	US-10-425-115-75767	Sequence 75767, A

ALIGNMENTS

RESULT 1

US-09-938-842A-1502
; Sequence 1502, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24

```
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1502
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1502

Alignment Scores:
Pred. No.: 5,91e-263 Length: 1977
Score: 2595.50 Matches: 498
Percent Similarity: 87.73% Conservative: 81
Best Local Similarity: 75.45% Mismatches: 77
Query Match: 74.69% Indels: 5
DB: 9 Gaps: 3

US-10-762-049-18 (1-680) x US-09-938-842A-1502 (1-1977)

QY 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
DB 1 ATGGGACGGAGGACTACATTCCTCAAGGAGCGGAGGAGCTTACCGCCGCCACCAC 60
QY 44 ---ValGluValProProGlnProPhePheLeuSerLeuLysTyrSerLeuGlu 62
DB 61 ACGGTGGAGGCTCCACAACCTCAACCGTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120
QY 63 ThrPhePheProAspAspProLeuArgGlnPheLeuAsnLysProAlaSerLysLysPhe 82
DB 121 ACTCTGTTCACAGCAGCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 180
QY 83 MetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLysTyrThrPheGln 102
DB 181 GTGTTAGGCTCAAACTTCCTCCGATTTTCGAATGGGCACCGCTACAAATCTCAAG 240
QY 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
DB 241 TTCTTCAAAATCAGATCTCATCGCGGGAATCACCATCGTAGTCTCGCCATCCCTCAGGCG 300
QY 123 IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIle 142
DB 301 ATCAGTTACGCCAAATCTGCTAACTTCGCCCCCAATTTCTGGCCCTTTATTGAGTTTGT 360
QY 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
DB 361 CCGCCATTTGGTATACCGGCTGCTAGGAGTTCAAGGAGCTTACCGGTGGACCGGTTGCG 420
QY 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
DB 421 GTTGGCTCTCTGTGACAGTGGGTGCTGAGCAAAAGAGTTGATGCTGAGAAAGATCCT 480
QY 183 LysLeuTyrLeuHisIleuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
DB 481 AAGCTTTTACCTTCACCTTGTCTTTCACCGCCACTTTTTCGCCGCGCTTCTCGAAGCCTCT 540
QY 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleG 222
DB 541 CTTGGAAATTTTCAAGGTTAGGGTTCATAGTGGACTTTTATCGCATGCAACGATAGTAGG 600
QY 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242
DB 601 ATTCATGGGAGGAGCAGCAGCGTGTGAGTCTGCAACAGCTTAAGGGTATTTTCGGACT 660
QY 242 uGluHisPheThrHisGlyAlaAspIleSerValMetArgSerValPheThrGlnTh 262
DB 661 TAAACATTTTACAGACTCTACCGATGTTATCTCTGTATCGGCTTCGTTTCTCCCAAC 720
QY 262 rHisGluTyrArgTyrPheGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
DB 721 TCACGAGTGGAGATGGGAAGTGGCGTTCTTGGAGTGGTTTCTATTCTTCTCTCTC 780

282 rThrArgTyrPheSerLysLysArgProArgPhePheThrValSerAlaMetAlaProLe 302
DB 781 CACCATGATTTTGAGCATCAGAAACCAAAATTTCTTTGGTGGCGGCGATGCTCCTTT 840
QY 302 uThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVa 322
DB 841 GACCTCAGTGATTTCTTGGAAAGTCTCTTGGTTTACTTTCACCTCAGCTGAGAGACATGGTGT 900
QY 322 lGluValIleGlyLeuLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342
DB 901 TCAAGTGATAGGGACCTGAAGAAAGGTTGAATCCACTCTCGGTTCTGATCATCTT 960
QY 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLe 362
DB 961 TACTTCCCCTTACATGTCACAGCTGTCAAAACTGGCCTCATCACTGGCATCATTTGCTCT 1020
QY 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGl 382
DB 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTTTCGGATGTTCAAGAACTCAACATAGACGG 1079
QY 382 yAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy 402
DB 1080 GAACAAAGAGATGATAGCGTTTGGATGATGAACATCGTTGGTCTCTTCACATCTTGTTA 1139
QY 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422
DB 1140 CCTCAACACCGGACCATTTTCAAGTTCGGCAGTGAACATACAACGGGGTTCGAAGACCGC 1199
QY 422 alaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLe 442
DB 1200 AATGTCCAACATAGTAGGCGATTCGGGTTATGTTCACACTCTCTCTCACACCGCT 1259
QY 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462
DB 1260 TTTTCTACTACACACACTCTCTCTCTGCCATCATCATATCGGCAATGCTCGGACT 1319
QY 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482
DB 1320 CATTCGACTATCAAGCTGCCATCCATCTCTGGAAGTTGACAAAGTTGCACTTCTCTGCTG 1379
QY 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI 502
DB 1380 CATGACGCGCTACGTGGGTCGTATTCGGCAGTGTAGAGATTGGACTCGTACGACGCT 1439
QY 502 eValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuGl 522
DB 1440 GCGGATATCTATAGCGAGGTTGTTCTGTTGTGTCGAGCCCAAAACTCGCGGTGAAGGG 1499
QY 522 yAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa 542
DB 1500 AAACATACCAACACGATGATCTATAGGAACACTGAGCAGTACCCATCATCAAGAACCCT 1559
QY 542 lProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrIle 562
DB 1560 TCCCTGGTATTCATCTTGGAGATTGATGCTCCCATCTACTTTGCTAATGTCAGTTACTT 1619
QY 562 uArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThrGlyGl 582
DB 1620 GCGTGAAGAATCATAAAGTGGATTGATGAAGAGAGAGAGAGCTTAACAATCAGGAGA 1679
QY 582 uThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGl 602
DB 1680 GAGCAGCTTACAATATATTATCTCGATATGTCAGCTGTTGGTAAATATATCGACACAAGCG 1739
QY 602 yIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeuValIle 622
DB 1740 TATTAGCATGATGTTGGAAATTAAGAAATGCTATTACAGGAGAGCGCTTAAAGTTGGTATT 1799
QY 622 uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641
DB 1800 GTCAATCCAAAGAGAGAGGCTCGTGAAGAAATTACACAGATCCAAATTCATCGGTGATCA 1859
QY 641 sLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAs 661
```

```
Db 1860 TTTGGGCAAGAGTGGATGTTCTTAACGGTAGGGAAGCAGTGGAGGCTTGTAGCTACAT 1919
Qy 661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 1920 GCTTCACACGCTTTAAACACCGACCG---GCCTCAAAAACGAGCGCTTGGAAACAACGTA 1974

RESULT 2
US-09-938-842A-1502
; Sequence 1502, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1502
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1502

Alignment Scores:
Pred. No.: 5,91e-263 Length: 1977
Score: 2595.50 Matches: 498
Percent Similarity: 87.73% Conservative: 81
Best Local Similarity: 75.45% Mismatches: 77
Query Match: 74.69% Indels: 5
DB: 11 Gaps: 3

US-10-762-049-18 (1-680) x US-09-938-842A-1502 (1-1977)
Qy 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
Db 1 ATGGGCACGGAGGACTACACATTCCTCAAGGAGCGGAGGCTTCACCGCCGCCAC 60
Qy 44 ---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 61 ACGGTGGAGGCTCCACAACTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTCAAGGAA 120
Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 121 ACTCTGTTTCAGACGACCTTTTAGCAATTTAAGAACCAAAATGATCAAGAAATTT 180
Qy 83 MetLeuGlyLeuGlnPhePheProPheGluTrpAlaProLysTyrThrPheGln 102
Db 181 GTGTAGGCTCAAAATCTCTCCCGATTTTCGAATGGCCACCGCTCAATCTCAAG 240
Qy 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 241 TTCTTCAAAATCAGATCTCATCGCGGAATCACCATCGCTAGTCTCGCCATCCCTCAGGC 300
Qy 123 IleSerTyrAlaLysLeuAlaAsnLeuProPheLeuGlyLeuTyrSerSerPheIle 142
Db 301 ATCAGTTACGCCAATCTGCTTAACCTGCCCCCAATCTTGCCCTTTATTCAGTTTGTGA 360
Qy 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 361 CCGCCATGTGTATACGGGTGCTAGGAGTTCGAAGGACTTAGCGGTGGGAACGGTTGCG 420
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
```

```
Db 421 GTTGGCTCTCTGTTGACAGTGCAGTGCAGCAAGAAGATTGATGCTGAGAAGAATCTCT 480
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAa 202
Db 481 AAGCTTTACTTACCTTCCACCTTGTTCACCGCCACTTTTTCGCGCGCGTCTCTCGAAGCTCT 540
Qy 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleLeuG 222
Db 541 CTTTGGAAATTTTCAGGGTTAGGGTTTCATAGTGGACTTTTCTATCATCATGCAACGATAGG 600
Qy 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242
Db 601 ATTATGGGAGGAGCAGCAGCGGTGTGAGTCTGCAACAGCTTAAGGTATTTTCGGACT 660
Qy 242 uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnTh 262
Db 661 TAAACATTTTCACAGACTCTACCGATGTATCTCTGTCTCATGCGTTCGTTTCTCCCAAC 720
Qy 262 rHisGluTrpArgTTPGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
Db 721 TCACGAGTGGAGATGGAAAGTGGCGTCTCTGGATGTGTTTCTATTTCTTCTCTCTC 780
Qy 282 rThrArgTyrPheSerLysLysArgProArgPhePheTTPValSerAlaMetAlaProLe 302
Db 781 CACGAGATTTTGAGCATCAAGAAACCAAAATCTTTTGGTGGCGCGGATGGCTCTCTT 840
Qy 302 uThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVa 322
Db 841 GACCTCAGTGATTTCTTGGAAAGTCTCTTGGTTACTTCACTCAACGCTGAGAGACATGGTGT 900
Qy 322 lGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342
Db 901 TCAAGTGTATAGGGGACCTGAAGAAAGGTTGATCCACTCTCCGGTTCGTATCTCATCTT 960
Qy 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLe 362
Db 961 TACTTCCCTTATCATGTCCACAGCTGTCAAAACTGGCCCTCATCTGGCATCATTTGCTCT 1020
Qy 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspG 382
Db 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTTGGCATGTTCAAGAACTACAACATAGACGG 1079
Qy 382 yAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy 402
Db 1080 GAACAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCCTTCAATCTTGTGA 1139
Qy 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422
Db 1140 CCTCAACACCGGACCACTTTCAAGGTCCGAGTGAACACTACAACCGGGTTGCAAGACCG 1199
Qy 422 aAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLe 442
Db 1200 AATGTCCAACATAGTATGCGGATTCGGGTATGTTTCACTCTCTCTCTCACACCGCT 1259
Qy 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462
Db 1260 TTTTCACTACACACCACTCGTCTCTCTGTCATCATCATATCCGCAATGTCGCACT 1319
Qy 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482
Db 1320 CATTTGACTATCAAGCTGCCATCTCTGGAAGTTGACAGTTGCAAGTTCTCTCTCTG 1379
Qy 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI 502
Db 1380 CATGACGCGCTAGTTGGGGTTCGTATTCGCGAGTGTAGAGATTGGACTCGTCGTAGCGGT 1439
Qy 502 eValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuG 522
Db 1440 GCGCATATCTATAGCGAGGTGTGTCTGTTTGTGTGTCGAGGCCAAAAAATCTCGCGTAGGG 1499
Qy 522 yAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa 542
```

```

1500 AAACATACCAAAACAGCATGATCTATATAGGAACACCTAGAGAGGTACCCATCATCAAGAACCGT 1555
Qy 542 lProGlyMetLeuIleLeuGluIleAspAlaProIleTy-PheAlaAsnAlaSerTyLe 562
1560 TCTGTGATATCTCATCTTGGAGATTGATGCTCCCATCTACTTTGCTAATGCCAGTTACTT 1619
Qy 562 uArgGluArgIleThrArgTrpIleAspGluGluGluGluArgIleLeysAlaThrGlyGl 582
1620 GCGTGAAAGAATCATAAAGTGGAATTGATGAAGAGCAAGAGAGAGCTTAAACAATCAGGAGA 1679
Qy 582 uThrSerLeuGlnTyrrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGl 602
1680 GAGCAGCTTCAATATATATATCTACGATATGTCTAGCTGTGGTGAATATCGACACAACGG 1739
Qy 602 YIleSerMetLeuGluGluValIleYsIleThrGluArgArgGluLeuGlnLeuValle 622
1740 TATTAGCATGATGCTGGAAATTAAGAAGTCATTGACAGAGAGCGCTTAAGTTGGTATT 1799
Qy 622 uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHi 641
1800 GTCAATCCAAAAGGAGAGGTGCGTGAAGAAATAACCCAGATCCAAATTCATCGGTGATCA 1859
Qy 641 sLeuGlyLysIleTrpIleTyLeuThrValGluGluAlaValGlyAlaCysAsnPheAs 661
1860 TTTGGGCAAAAGAGTGATGTCTTTAACGCTAGGAGAAGCAGTGGAGGCTTGTAGCTACAT 1919
Qy 661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnVal 680
1920 GCTTCACAGCTTTAAACCGAACCG---GCCTCCAAAAGAGCGCTTGGAAACACGTA 1974

RESULT 3
US-10-437-963-38184
; Sequence 38184, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 38184
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41845C.1
US-10-437-963-38184

Alignment Scores:
Pred. No.: 7,65e-235 Length: 2372
Score: 2330.50 Matches: 441
Percent Similarity: 83.84% Conservative: 83
Best Local Similarity: 70.56% Mismatches: 100
Query Match: 67.06% Indels: 1
DB: 19 Gaps: 1

US-10-762-049-18 (1-680) x US-10-437-963-38184 (1-2372)

Qy 43 GlnValGluValProProGlnProPhePheLysSerLeuLysTyrrSerLeuLysGlu 62
Db 149 AGGTGCGCATGCCGCGCGGAGACCCGTTCTTGAGACGCTGGGGGGGGAACATGAAGAG 208
Qy 63 ThrPhePheProAspProLeuArgGlnPhe---LysAsnLysProAlaSerLysLys 81

```



```

Db      1623 GCATGATGGGAGCATTTTGGAGTCATATTTTCATCGTGAGATATGCTTGCATTT 1682
Qy      501 AlaileValileSerValLeuArgValLeuLeuPheilleAlaArgProArgThrPheVal 520
Db      1683 GCGTTGCAATATCTCTTGTAATAATCTCTCAAGTACACGGCCAGAACAGTTTAA 1742
Qy      521 LeuGlyAsnilleProAsnSerValilleTyrArgAsnValGluHistiTyrGlnAsnAlaLys 540
Db      1743 CTTGGAAACCTTCACGAAACACTATATACAGGAATGTGAACAGATATCTGTATGTACC 1802
Qy      541 HisValProGlyMetLeuileLeuGluileAspAlaProIleTyrPheAlaAsnAlaSer 560
Db      1803 AAGTTTCCAGGGTGCTGATGTTAGAGTGAGTCTCAGCTATATACCTTACAAACTCTAAC 1862
Qy      561 TyrLeuArgGluArgIleThrArgTTPilleAspGluGluGluArgIleLysAlaThr 580
Db      1863 TATGTTAAAGAGAGAACTCTGAGTGGCTAAGAGATGAGGAGGAGCAACACAGGACCAG 1922
Qy      581 GlyGluThrSerLeuGlnTyrValilleIleAspMetSerAlaValGlyAsnilleAspThr 600
Db      1923 AAGTTTAAACAAAACCTGAGTTTCTAAATGTTGACCTGCTCTCTGTAATTGATATCGACACA 1982
Qy      601 SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db      1983 AGTGGAAATCATGCTTTGGAGGAGTTGGCGAAAGCTCTTGAAAAACGCAAAATTCAGCTG 2042
Qy      621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db      2043 GTTCTTACCAATCCCGGGCGGGGTGATCCAGAGCTCGCTCAGCGAAATTCACGGAC 2102
Qy      641 HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPhe 660
Db      2103 ATGATTTGGTGAAGCAACATATTTCTCAGCGTCGGCGACGCTGTG----- 2147
Qy      661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsn 678
Db      2148 -----ANGAAATTTGCTCTTAAGATGGTCGATGAATGCTGTGATGGAGT 2189

RESULT 6 15-182207
US-10-425-115-182207
; Sequence 182207, Application US/10425115
; Publication No. US2004021427A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 182207
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97750C.1
US-10-425-115-182207

Alignment Scores:
Pred. No.: 1-5e-183 Length: 2651
Score: 1845.50 Matches: 355
Percent Similarity: 74.41% Conservative: 116
Best Local Similarity: 56.08% Mismatches: 157
Query Match: 53.11% Indels: 5
DB: 20 Gaps: 3

US-10-762-049-18 (1-680) x US-10-425-115-182207 (1-2651)

Qy      41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60

```

```

Db      518 GTGCAACAGGTGGCGCGCGCGCGCGAGCACGCGCAGCAAGATGAAGGTGAGGGTG 577
Qy      61 LysGluThrPhePheProAspPheProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db      578 AAGGAGACCTTCTTCCCGACGACCGTTCGGGCGTTCAAGGGGCGACCGCGGGGACG 637
Qy      81 LysPheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThr 100
Db      638 CAGTGCTCATCGCGGTGAGTACCTCTTCCGCATCTCGACTGGGTGGCGAGTACTCC 697
Qy      101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
Db      698 TTGTGCTCTTCAAGTCCGACCTCGTCCGCGGCTCACCATTGCCAGCTTCGCATTCCT 757
Qy      121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140
Db      758 CAGGCGATTAGCTACGCGAAGCTGGCAAGCTTGCCTCCCATTAATCGGCTGATTTCGAGC 817
Qy      141 PheIleProProIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db      818 TTCGTGCCCGCATGGTGTACGGTGTGGGAGCTCCCGTACCTGGCGGTGGGCCCG 877
Qy      161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
Db      878 GTGTGATCTCGTGTGATCATGGGTCCATGCTGGCGAGCGCTGAGCCCCACCGCG 937
Qy      181 AspProLysLeuTyrIleHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
Db      938 GAGCCGACGCTGTCTCTGCGAGCTGGCTTACCTCCACCTGTTCGCGGGCTGGTGAG 997
Qy      201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
Db      998 GCCTCCCTGGGCATCTCAGGCTCGGCTTCGTCATCGACTTCCTGTCTCCAAAGGCGAGCTG 1057
Qy      221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeu 240
Db      1058 GTGGGGTTCATGGCGGCGCGCCCATCATCTGGCGGCTCGAGCAGCTCAAGGGCGTGTG 1117
Qy      241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
Db      1118 GGCATCGTCCACTTCACCAACCGAGATGGGCATCGTCCAGTCATGGCTCCGCTTCCAC 1177
Qy      261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
Db      1178 CACACCGAGTGTGTGTGGCAGCATCTCATGGGCTCTGCTTCTCTCTCTCTCTCTG 1237
Qy      281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
Db      1238 CTGTGGCGAGGATGTGAGCATCAGATGGCCAAAGCTTTTCTGGGTTCGGCGTGGCGG 1297
Qy      301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db      1298 CCCCTGGCATCGTCCACCATCTCGACGCTGCTGTTTCTCTCTTCTTCTTCTTCTTCT 1357
Qy      321 GlyValGluValIleGlyLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
Db      1358 GGCATCAGCATCATTTGGGCGAGCTCAAGTCGGCGCTGAATCGCCCTCTGCTGGGCAAGACTC 1417
Qy      341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
Db      1418 CTGTTTGACACGGCGTATTTAGGCTTCACCATGAGACTGGGCTTGTACCCGGAATCATC 1477
Qy      361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
Db      1478 TCCTGACCGAAGGAATAGCGTTGGTGAACATTTGCTCTCACTCAAGGACTACAGATA 1537
Qy      381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
Db      1538 GATGGAACACAGGAGATGATGCGCATAGGTTGATGAATGTTGTTGGTCTCTGCACATCA 1597
Qy      401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420

```



```
Db 1598 TGCTACCTAACACAGGTGGTCTCCCGCTCTGCTGTAAACACACACGCCGGCTCAAG 1657
Qy 421 ThrAlaAlaSerAsnIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db 1658 ACTGCCATGTCCACGTGATCATGGCGCTGACTGTGATGTCAGGTGCTTCTCATG 1717
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db 1718 CCACCTGTTCGTGTACACACCAACGTTGTCTCGGAGCGATCATCATCGCGGGTGATC 1777
Qy 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 1778 GGCTGATCATGATTCCCGGGGTGTACCATCTGGAAGATGACAAAGATGGATTTCTG 1837
Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db 1838 GTGTGGTTTGGCGGTTTGGCGGTCATCTTCATCTCAGTCCAAAGAGCCCTTGGCATA 1897
Qy 501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520
Db 1898 GCGGTTGGTATATCTATATTAGGGTGTGTGATGCAGATCAGGAGGCCCAAGATGGT 1957
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnIleAlaLys 540
Db 1958 CAAGGGAACATCAAGGGGACTGATATTACAGAGCTGCATCACTACAAAGGAGGCCAA 2017
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaIleAlaSer 560
Db 2018 AGAGTTTCTGGGTTCTTGATCTTGGCCATTGAAGCAGCCGATAAATCTGCCAACTCCA 2077
Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db 2078 TACCTGAATGAAGAGTAAAGATGGATA-----GAGGAAGATCTTTTGAACAGGAT 2131
Qy 581 GlyLeuThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db 2132 AAACATACAGAACTCCATTTCATAATCTTGGATCTGTGAGCTGTCCTGCAATTGACACA 2191
Qy 601 SerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db 2192 AGTGGCATAGCGTTCCTCATTTGACATAAAGAATCAATAGAGAAACGTGGTCTGGAGCTT 2251
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLys---SerLysPheGln 639
Db 2252 GTGCTTGTCAATCCAACTGGAGAAGTCATGGAGAAAATACACGTGCAAAACGAGGCTGAA 2311
Qy 640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db 2312 AACTATTTTAGCCAGATTGCTTGTATCTGACCCTGCGGCAAGCAATC-----GCTTCA 2365
Qy 660 PheAsnLeuArgAlaSerLysThrAsnProLysLysAsp 672
Db 2366 CTTTCTGCATTCGCAAGATGACAAAACCCCTAAATGGAT 2404

RESULT 7
US-10-424-599-65103
; Sequence 65103, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 65103
; LENGTH: 2366
; TYPE: DNA
; ORGANISM: Glycine max
```

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_29800C.1

US-10-424-599-65103

Alignment Scores:

Pred. No.:	1-59e-183	Length:	2366
Score:	1844.50	Matches:	367
Percent Similarity:	73.64%	Conservative:	133
Best Local Similarity:	54.05%	Mismatches:	165
Query Match:	53.08%	Indels:	15
DB:	18	Gaps:	5

US-10-762-049-18 (1-680) x US-10-424-599-65103 (1-2366)

Qy 8 SerTyrIleThrHisIleCysLeuLeuArgAsnThrIleIleGluAspMetGlySerVal 27

Db 73 ACTTACATCACACCACTTTCTCTTATA-AGTCTGTTCTCTTCAAAATGGAGACTAAT 131

Qy 28 Asp-----TyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43

Db 132 AATGCATGTACTATGCAATTCATTTCATAGATAGATGTCC-----ATGGAAGTGCAACAA 185

Qy 44 ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThr 63

Db 186 GTTGTGCCACCACTTCACAAGAGCACCTTCAGAACTCAAGGGTAGGCTCAAGGAAACT 245

Qy 64 PhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPheMet 83

Db 246 TTCCTCCCTGATGATCTCTGCGCCNATTCAGGGACACACCTCTTAAGAGAAAACGTGATC 305

Qy 84 LeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThrPheGlnPhe 103

Db 306 CTTGGAGCTCAATATGTGTCTTCTTCCAAATGGGGTCTTAAGTATATAATCTCAAACTC 365

Qy 104 LeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIle 123

Db 366 TTCAAATCTGACCTTGTCTGCGCTCACTATTGTAGCTGGACATCCCGCAGGGAATG 425

Qy 124 SerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIlePro 143

Db 426 AGTTATGCTTAGCTTGCAGAGTCTTCTCCAAATTTTAGGACTTTATTCTAGTTTGTCCA 485

Qy 144 ProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaVal 163

Db 486 CCACCTGTCTATGCTGTCTTGGAAAGCTCAAAGGACCTTGACAGTTGGACCTGTTCTATT 545

Qy 164 GlySerLeuMetCysSerMetLeuSerAsnAlaValAspProAsnGluAspProLys 183

Db 546 GCTTCTTGTGATGGGATCCATGTTGCATCAGGAAGTGTCTCCCAACACACATCAAT 605

Qy 184 LeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeu 203

Db 606 CTGTTTCTTACGTAGCTTTCATTCACATTTATTGGCGGTCTCTTTCAAGCTTGTCTT 665

Qy 204 GlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPhe 223

Db 666 GGAATCCTAAGGCTAGGCTTCATAATGATTTCTTCTTAAGGCCATTCTTATTGGGTTTC 725

Qy 224 MetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeuGlyLeuGlu 243

Db 726 ATGGCTGGAGCTGCTATTATTGCTCTCATCTGCAACAGCTCAAGAGCTGCTGGAATCACA 785

Qy 244 HisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHis 263

Db 786 CATTTCACTAATCAGATGGGTCTGATTCCTGTTATGACTTCGTGTTTTTTCACAAATACAT 845

Qy 264 GluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuSerThr 283

Db 846 GAGTGGTCATGGCAACCAATATTGATGGGATTTGCTTCTTGGTACTACTACTATTAGCA 905

Qy 284 ArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeuThr 303

Db 906 AGACACGTTAGCATTAAGGAAACCAAACTATTCTGGGTCTCAGCTGGAGCTCCTCTTATG 965

Qy	304	SerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyValGlu	323
Db	966	TGTGTCAATCATCTCTACCTCTTGGTCTTTTGGCAATTAAGGCTCAAAATCATGGCATCAGT	1025
Qy	324	ValIleGlyGluLeuLysGlyLeuAsnProSerLeuThrAsnLeuValPheVal	343
Db	1026	CGGATTGGAAATTCGAACAGGAATAATCCTCATCATGGAAATATGTGTGCTCTTTTCAT	1085
Qy	344	SerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAla	363
Db	1086	GGAAAGTCACCTAGGCCTAGTTATGAAACACAGGCTTATCACCGGGATTTGTCCCTTAACA	1145
Qy	364	-GluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAs	383
Db	1146	GGAAAGGATTTCAGTAGGAAGACATTTGCAAGCTCTCAAAACATCAAAAGTGGATGGAAA	1205
Qy	383	nLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLe	403
Db	1206	TAAGAAATGATGGCAATGGGTTTATGAATGTGTGGCTCTCTCACTTCTCTGCTAATGT	1265
Qy	403	uThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAl	423
Db	1266	TACAAACAGGTCTTCTCTCGTCTCAGCAGTTTAAACAATGCAGCGCGCAAAACAGCTGT	1325
Qy	423	aSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPh	443
Db	1326	GTCAAAATGTAGTGATGTCTGTGACAGCATGTGTGACATCTCTTTTCTCATGCCATTTGTT	1385
Qy	443	eHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIle	463
Db	1386	TCAAATACAGCTTAATGTCTGTCTGGCGCAATCATAGTCACAGCAATATGTGCGCTCAT	1445
Qy	463	eAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMe	483
Db	1446	CGATCTCCCGCTGCTTGAACATTTGGAAGATCGACAAATTCGATTTTGTGTGATGAT	1505
Qy	483	tSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleVa	503
Db	1506	GACTGCTTTCTTAGGTGTCTTTTCATCTCTGTCCAGGAGGCCCTCTCTGTGCTGTGG	1565
Qy	503	lIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAs	523
Db	1566	ATTATCACTTGAAGTACTCTTTGCCAAATTACAGACCGGAAACACAGTGATGTGGGGAA	1625
Qy	523	nIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValPr	543
Db	1626	GATACCAAGGAACACACATATATAGAAATCTTGATCAATACAAGGAAGCTGTGAGAATACC	1685
Qy	543	oGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuAr	563
Db	1686	TGGATTTCTTATTTTAAGCATTTAGGCTCCCATCAATTTTGCACCAATCATCATCTCAA	1745
Qy	563	gGluArgIleThrArgTrrIleAspGluGluGluArgIleLysAlaThrGlyGluThr	583
Db	1746	CGAGAGAACGTTAGATGGATTGAAGAGAGAGAGACACAAATAAGGAA-----CAATT	1799
Qy	583	rSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIl	603
Db	1800	AAGCCTTCGATTCCTAGTATTGGAAATGTTCAGCTGTGAGTGCTGTGTGACACAAAGTGAAT	1859
Qy	603	eSerMetLeuGluGluValLysLysIleThrGluArgGluLeuGlnLeuValLeuVa	623
Db	1860	CTCACTTTTCAAGGAGTTGAAGCAACACTGGAAAGAAAGGTTGTGAGCTGTGTGGT	1919
Qy	623	lAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn---HisLe	642
Db	1920	CAATCTCTGCTGAGTTCATAGAAAAGCTTTAAAAAAGCAGATGAAGCTAATGATTTTCAT	1979
Qy	642	uGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLe	662
Db	1980	AAGAGCAGATAAACCTTTTCTTGAAGTTGGAGAGCGCTGTAGCT-----TCATCT	2027

```

QY 662 uA9GalaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 2028 TTCTTCAGCAATGAAAGGCCAATCATCAACCATTTACAGAAGGGGCACACACAATT 2082

RESULT 8
US-10-437-95853
; Sequence.95853, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 95853
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94005C.1
US-10-437-963-95853

Alignment Scores:
Pred. No.: 3,69e-180 Length: 2001
Score: 1811.50 Matches: 340
Percent Similarity: 73.10% Conservative: 122
Best Local Similarity: 53.80% Mismatches: 167
Query Match: 52.13% Indels: 3
DB: 19 Gaps: 3

US-10-762-049-18 (1-680) x US-10-437-963-95853 (1-2001)

QY 43 GlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 106 GAGGTGAACCTGTGCGGGCGCGCCCGTTCGGGAGAACCTGTGTGCGACCTCGCGGAG 165

QY 63 ThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 166 ACGTTCTTCCCGACGACCCCGTTCGCGGGTTTCGGCGCGCTGCCCGCGCAGCGCGCG 225

QY 83 MetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThrPheGln 102
Db 226 TGGTGC CGCGGTCAAGTACTTCGTC CGCGCTGGGACTGGGTCCCGGCTACGCGCTCGAC 285

QY 103 PheLeuLysAlaAspLeuIleAglYileThrIleAlaSerLeuAlaIleProGlnGly 122
Db 286 AAGTTCAGTTTCGACCTCTCTCGCGCGCATCACCATCGCAGGCTCGCCATCCGCGAGGC 345

QY 123 IleSerTyraLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIle 142
Db 346 ATCAGTACGCGCGCTCGCAACCTCCCTCCAATCATCGGCTCTATTTCGAGCTTCGTG 405

QY 143 ProProLeuIleTyraLamMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 406 CCGCGGTGATGTACGCGGTGTTCGGGAGCTCCAAACACCTCGCGGTGGGACGGTGGCG 465

QY 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValaAspProAsnGluAspPro 182
Db 466 CGCGGCTCGCTGCTGCTGCTGCATCATCGAGACGGAGTGGCGCGGACGNAACCCG 525

QY 183 LysLeuTyraLysLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 526 CAGCTGTACCTGCAGCTCTTCTACACCGCGCGCTTCTTCACCGGCTCTTCAGACCGCG 585

```



```
Db 613 GACTGGGGCGGCACATACCTTCGCAAGTTGAGGGGCGACCTGCTCCGGCCCTCAC 672
Qy 114 IleAlaSerLeuAlaIleProGInGlyIleSerTyrAlaIlySerLeuAlaSerLeuPro 133
Db 673 ATTGGCAGCCTCTGCATACCCAGGACATCGGTATATGCGAAGCTTGCTGGCCCTATTACCA 732
Qy 134 IleLeuGlyLeuTyrSerSerPheIleProLeuIleTyrAlaMetMetGlySerSer 153
Db 733 AATTATGAGCTGTACAGCAGCTGTGATACCGCCTTGATATACGCGATGATGGGAAGTTCC 792
Qy 154 ArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMetLeuSer 173
Db 793 AGGATATAGCCATTGTGTCAGTGGCGCTGTTTGGCTGCTGTGGCACTCTCCCTCCAG 852
Qy 174 AsnAlaValAspProAsnGInAspProLeuLeuTyrLeuHisLeuAlaPheThrAlaThr 193
Db 853 AATGAGTTTGCATCCCAAGAAGATCAGGAAGAGTACACCCGGCTGGCCCTTCACTGCAACG 912
Qy 194 LeuPheAlaGlyValPheGInAlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAsp 213
Db 913 TTCTTGGCTGGGTTCACCCAGGAGTGTCTCGGATTCCTCAGGCTAGGGTTTATCATAGAG 972
Qy 214 PheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValValCysLeu 233
Db 973 TTCTTGGCCATGCTGCCATCGTTGGATTTCATGGCGGAGCGGCCATCACCATTGGCCCTT 1032
Qy 234 GInGlnLeuIlySerIleLeuGlyLeuGluHisPheThrHisGlyAlaAspIleIleSer 253
Db 1033 CAGCAGCTTAAGGCTTCCTTGGATTTGCAAACTTCACCAAGAAGACTGACATCATCTCG 1092
Qy 254 ValMetArgSerValPheThrGInThrHisGlu---TrpArgTrpGluSerAlaValLeu 272
Db 1093 GTCATGAAATCAGTCTGGGGAATTTTCACCATGGGTGGAACTGGCAGACGATATTGATC 1152
Qy 273 GlyCysValPheIlePheLeuLeuSerThrArgTyrPheSerIlySerIlyArgProArg 292
Db 1153 GGAGCATCTCTCGGCATTCTCCCTGGTTGCCAAATATACATTGCCAAGAAGAAATAAGAA 1212
Qy 293 PhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuVal 312
Db 1213 CTCTTTTGGTGGCTGCAATTGACCACTCACTTCGGTGATCATATCCACCCCTTTTGTG 1272
Qy 313 TyrPheThrHisAlaGluIlyHisGlyValGluValIleGlyLeuLeuIlySerGlyLeu 332
Db 1273 TATATCACTCGAGCAGACAGCATGGTGTCTGATGTCAGTACATACAAAGAAAGGCATC 1332
Qy 333 AsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLys 352
Db 1333 AATCCACCTTCAGCTAGTCTCATATATCTTCAGTGGGCCCAACCTGATGAAGGATTTCAGA 1392
Qy 353 ThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPhe 372
Db 1393 ATAGGAGTATAGCTGCATGATAGCGCTAACAGAGCAATTCATAGGAAGACATTC 1452
Qy 373 AlaMetTyrIlyAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMet 392
Db 1453 GCTGGCCGTAAGGACATCAAGATTGACGGAAACAAGGAATGTTGCTCTAGGAACAATG 1512
Qy 393 AsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAla 412
Db 1513 AACATTGTGTGTTCAATGATCTTCTGTACGTAGCTACAGGTTCTTTCTCGCGGTGAGCA 1572
Qy 413 ValAsnTyrAsnAlaGlyCysIlyThrAlaAlaSerAsnIleIleMetSerLeuAlaVal 432
Db 1573 GTCAATTACATGGCTGGCTGCCAAACAGCGGTATCCATATTGTTATGTCAATTGTAGTA 1632
Qy 433 MetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSer 452
Db 1633 CTGCTCACATGGAGCTGATTACTTCCATTGTTCAAGTATACGCCCAATGCCATCTTCT 1692
Qy 453 AlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaIleHisLeuPhe 472
```

```
Db 1693 TCATCATCATATACAGCAGTGTGCTGCTAGTCAGTACACACAGCATACCTTATCTGG 1752
Qy 473 LysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPheGly 492
Db 1753 AAGTTGATAAACTTGACITTCCTGGCATGCTTAGGAGCATCTTTTGGCGTCATATTTCA 1812
Qy 493 SerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuPhe 512
Db 1813 TCAGTGGAGTATGCTGCTAATTTGGTGGATCTCACTCGCTAAAATTTCTTACAA 1872
Qy 513 IleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrAspAsn 532
Db 1873 GTTACAGCAGCCAGAACAGTTTGTGTGTACCTTCCAGAACACATATATATAGGAAT 1932
Qy 533 ValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAla 552
Db 1933 ATCGACCACTATCCAGAAGCTACCTTGGTACAGGAGTGTGTAATAGTTAGGGTGGACTCA 1992
Qy 553 ProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGlu 572
Db 1993 GCTATCTACTTCAAACTCGAACTATGTTAAAGACAGAAATCCTGAGGTGGCTGAGGGAC 2052
Qy 573 GluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMet 592
Db 2053 GAGGAAGAAGCTCAGCAGGAACAGAGCTACAGAAACTGAGTTTCTGATTTGTAGCTA 2112
Qy 593 SerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLysLys 612
Db 2113 TCTCCGGTAAATGACATTCATCAAGTGAATCCATCCCTTGGAGGATCTATTTAGAGCA 2172
Qy 613 ThrGluArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLysLys 632
Db 2173 CTTGAAAACGCAAAATCCAGCTGATTCTAGCCAATCCGGGGCCAGCTGTGATCCTGAAG 2232
Qy 633 LeuAsnLysSerIlySerPheGlnAsnHisLeuGlyLysLysTrpIleTyrLeuThrValGlu 652
Db 2233 CTCGATCAGCAAAATTCAGGATCTATCGGTGAAGACAGATATTCCTGACTGTGCGG 2292
Qy 653 GluAlaVal 655
Db 2293 GACGCAAGTG 2301
RESULT 10
US-10-425-114-15719
; Sequence 15719, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15719
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-059-F10_FLI
US-10-425-114-15719
Alignment Scores:
Pred. No.: 9.87e-175 Length: 2656
Score: 1762.00 Matches: 334
Percent Similarity: 73.23% Conservative: 120
Best Local Similarity: 53.87% Mismatches: 162
Query Match: 50.71% Indels: 4
```

```
DB: 18 Gaps: 2
US-10-762-049-18 (1-680) x US-10-425-114-15719 (1-2656)
QY 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
DB 299 CTCACAAGGTGTCTTCGCGGAGCGCGGACGACGCGGAAGCGTGCAGCCGCTG 358
QY 61 LysGluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
DB 359 GCCAGAGTGTCTTCGCGAGCACCGCTGCACAGTTCNAGAACACAGTCGTCGGCGCG 418
QY 81 LysPheMetLeuGlyLeuGlnPhePheProPheProLeuTrpAlaProLysTyrThr 100
DB 419 CGCTCGTGTGGCGTGCACACTCTTCTCCCATCTTCACAGTGGGGTCCGCGCTACAGC 478
QY 101 PheGlnPheLysAlaAspLeuLeuAlaGlyLeuThrLeuAlaSerLeuAlaLeuPro 120
DB 479 CCAGCGCTCTCGGCTCCGACCTCGTCGCGGCTCACCATTGCCAGCTCGCCATCCG 538
QY 121 GlnGlyLeuSerTyrAlaLysLeuAlaAsnLeuProProLeuGlyLeuTyrSerSer 140
DB 539 CAGGGAATCAGTCACGCAAGCTCGCAACCTCCGCAATGCTTGGCTATATCCAGC 598
QY 141 PheLeuProLeuLeuTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
DB 599 TTCGTGCGCGCTCATCTACGCGCTGCTGGGAGCTCGCGGACCTGGCGGTGGGCGC 658
QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
DB 659 GTGTCCATCGCGTGTGTGTGATGGGGTCCATGCTCCGGAGCCGCTGCGCGGACGAG 718
QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
DB 719 CAGCCGCTCTTACTGTCAGCTGGGCTTCACGCCACCTTCTCCGCGGCTTCCAG 778
QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuLeuValAspPheLeuSerHisAlaThrIle 220
DB 779 GCGTCCCTGGGATTCCTCAGGCTGGGCTTCATGCTGAGCTTCCTGTCGAAGGCGAGCTG 838
QY 221 IleGlyPheMetGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeu 240
DB 839 ACGGGCTTCATGGCGCGCGCGCTCATGCTGCTCGCAGCAGCTCAAGGGCTGCTC 898
QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleSerValMetArgSerValPheThr 260
DB 899 GGATCTCCCACTTACCTCCACATGGGATTCCTCGACGTATGCGCTCCGCTGCTCAAC 958
QY 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280
DB 959 CGCCACGACGAGTGAAGTGGCAGCAGTCTGTCATGGCTCCGCTTCCTCGCCATCTC 1018
QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
DB 1019 CTCTCTACGCGCCAAATCAGCGCCAGGAACCCCAAGCTTTTCTGGGTATCAGCAGGTGCT 1078
QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGlyLysHis 320
DB 1079 CCCTGGCGTGGTGATATCTCCACATCTCTCTTCATCTGG-----AAATCCCC 1132
QY 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
DB 1133 AGCATCAGTGTATTGGCATCTCTCCAGGGAGTGAACCTCTCTTCGCGAACATGCTC 1192
QY 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyLeuValValGlyIleIle 360
DB 1193 AGCTTCAGCGGCTCTATGTGGCGCTGCACATCAAAACCGGGATCATGACAGCATCTG 1252
QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
DB 1253 TCTTTAACAGAGGATCGCAGTGGGAGGACCTTCGCTGCCATCAACACTACCAAGGTG 1312
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
```

```
DB 1313 GACCGGAACAAGAGATGATGCGCATCGGCTGATGAATGAACATGCGGGCTCTCTCGCCTCC 1372
QY 401 CysTyrLeuThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
DB 1373 TGCTACGTGACACAGCGGGTCTTCTCCCGGTTCGGGGTGAATACAGCGGGCTGCAGG 1432
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
DB 1433 ACGGCGCTCTCAACGTCGTGATGCGGCGGGTGTGGTGCAGCTGCTGTCTCTCATG 1492
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
DB 1493 CGGCTGTTCCACTACACCCGAACTGATCTCTGGGGCGCATCATCATCAGCGGTGGTG 1552
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
DB 1553 GGGCTGGTGGAGCTGCGGCGCGCCCGAGGCTGTGAAGGTGGACAAGCTGACATTCCTG 1612
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGlyLeuIleGlyLeuValIle 500
DB 1613 GGTGCGTGGCGGCTTCTCGCGTGTGTGTGTGTCGTCGTCGACAGCGGCTGGCGCTC 1672
QY 501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520
DB 1673 GCGCTCGGCATCTCGCTCTCAAGTCTCTGTCAGGTCAACCGCCCAACGTCGTGGTG 1732
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
DB 1733 GAGGCGCTCGTCCCGGGACGACAGACTACCGCAGCGTGGCGCAGTACCAGCGGCGCTC 1792
QY 541 HisValProGlyMetLeuLeuLeuGlyLeuIleAspAlaProIleTyrPheAlaAsnAlaSer 560
DB 1793 CCGCTGCGGGCTTCTCGTGTGCGGCTCGAGTCCGCGCTCTACTTCGCCAATCCATG 1852
QY 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluGluArgIleLysAlaThr 580
DB 1853 TACCTGGTGGAGCGGCTCATGCGCTACCTCCCGCAGCAGGAGGCGCGCTCAAGTCC 1912
QY 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
DB 1913 AACACCCCTCCATCCGATGCGTCTCGACATGGGCGCGCTCGCGCGATCGACAG 1972
QY 601 SerGlyIleSerMetLeuGluGluValLysIleThrGluArgGluLeuGlnLeu 620
DB 1973 ACGGCTTAGACGCGCTGTCGAGCTCAAGAAAGTCTTGACAAAAGAAACATCGAGCTG 2032
QY 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
DB 2033 GTGCTTGCCACCCGCTGGGTGCGTGGCGGAGAGATGTTCAACTCGCGCGTGGCGCAG 2092
QY 641 HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaVal-----GlyAlaCys 658
DB 2093 AGCTTCGGTTCGGCGCGCTCTTCTTCAGGTAGCGGAGGCGCTCGCGCGGGGCGTGC 2152

RESULT 11
US-10-425-115-75763
; Sequence 75763, Application US/10425115
; Publication No. US2004021472A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plasmids
; FILE REFERENCE: 38-21(53222) B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 75763
; LENGTH: 2713
; TYPE: DNA
```

; APPLICANT: Kovalic, David K.

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 73410
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73694C.1
; US-10-437-963-73410

Alignment Scores:
Pred. No.: 6.84e-172 Length: 2429
Score: 1734.50 Matches: 333
Percent Similarity: 70.93% Conservative: 116
Best Local Similarity: 52.61% Mismatches: 163
Query Match: 49.91% Indels: 21
DB: 19 Gaps: 3

US-10-762-049-18 (1-680) x US-10-437-963-73410 (1-2429)

QY 41 ValHieGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2261 CTCACAAAGGTGTCGGTGGCGGAGAGGAGGTGCGACGCGAAGCGCTGAGCGCGGCTG 2202

QY 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2201 GCGGAGGTGTTCTTCCCGGACGACCGCGCTCCACAGTTCAAGAACACAGCTGCTCCGCGCG 2142

QY 81 LysPheMetLeuGlyLeuGlnPhePheProilePheGluTrpAlaProLysTyrThr 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2141 CGGCTGGTGTCCGGTGGCGAGTACTTCTTCCCATCTTTCATTTGGGGTCCGAGCTACAGC 2082

QY 101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2081 CTCGCGCTCTCGCTCCGAC----- 2061

QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSer 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2060 ---GGAATCAGCTACGCGCAAGCTCGCAACCTCGCTCCAATCATTTGAGCTATATTCGAGC 2004

QY 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2003 TTGTCGCGCGCTGATCTACTCGTTGTTGGTAGCTCGCGGAGCTGCGTGTAGGGCGCG 1944

QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValaAspProAsnGlu 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1943 GTCTCGATAGCGTCGTGGTGATGGGTGCGATCTCGGCGAGCGGTGTGCGCGGACCAG 1884

QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1883 GAGCCCATCTCTACCTCCAGCTCGCTTCACCTCCACCTTCCTCCGCGCGCTTCCTCAA 1824

QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValaAspPheLeuSerHisAlaThrIle 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1823 GCCTCCCTCGGCTTCTCCGGCTGGGGTTCATCGTGGATTCTTGTGGAAGGCGACGCTG 1764

QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1763 ACGGATTCATGGCGCGCGCGGATCATCGTGTGCTGCGAGCAGCTCAAGGGATTGCTC 1704

QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1703 GGGATCATCATTTCACTCGCAGATGGGGTTCGTCAGGTGATGCACCTCCGCTTCAAG 1644
```

```
QY 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1643 CACCACGAGTGGCGTGGCAGACCCTCTCATGGGGGTCCCTTCTCTCCCGCTCTC 1584

QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1583 CTCACACAGCCCATCATCAGCGCCAGGAACCCAAAGCTTTCTGGGTGTGACAGCTGCT 1524

QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1523 CCACGTGACATCAGTGCATCTCTACCATCTCTCGTTCGTTCAGCAAGCT-----CAT 1470

QY 321 GlyValGluValIleGlyLeuLysGlyLeuAsnProPhePheSerLeuThrAsnLeu 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1469 GGCATCAGTGTATTGGCGATCTCCCAAGGATTTGAACCTCTTCACGGAATATGCTG 1410

QY 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1409 ACCTTCAGTGGCTCTACGTAGGACTGGCTTTAAACACACAGGAATTATGACTGGCATACTA 1350

QY 361 SerLeuAlaGluGlyIleAlaValAlaGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1349 TCTCTCACTGAAGGATAGCAGTAGCAGGACATTTGCATCCATCAACAACCTACCAAGTT 1290

QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValIcLysSerPheThrSer 400
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1289 GATGGAACAAGGAGATGATGGCCATTGGTGTCTAGAACATGGCTGGCTCATGTGCCTTC 1230

QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1229 TGTATGTCACACAGGATCGTTCCTCGAGTCCGCGGTGAACCTACAGCGCGGGTGAAG 1170

QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1169 ACGGCGGTGTCAAACATCGTGTGCGTGGCGGTGTGGTACGCTGCTGTTCCTGATG 1110

QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1109 CGGCTGTTCCATCACACCCCAACGTGATCTCTCGCGATCATCATCACCGCGGTGATC 1050

QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1049 GGCCTCATCGAGTCCGGGCGCGCCAGCGTGTGGAAGTTCGACAAGCTCGACTTCCTC 990

QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 989 GCGTGCATGGCGCTTCTCGCGCTGCTCTCTCTCGTCCAGATGGGCTTCGCCATC 930

QY 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 929 GCGTGGCATCTCTCTTCAAGATCTCTCCAGGTCAACCGGCCAACATGGTGGTGC 870

QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 869 AAGGGCGTCTCCCGCACGCGAGTACCAGCATCGCGCAGTGGCGCAGTACAGGAGGCGCATG 810

QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 809 CGTGTGCGCTCTCTCTCGTGGGTGAGTCCGCCATCTACTTCGCCAATCTCATG 750

QY 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 749 TACCTCGCGAGAGGATCATGAGGTTCTCCGGGAGGAAGACAGCGGCGGCCCAAGTGC 690

QY 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValIcLysAsnIleAspThr 600
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 689 AACAGTGGCTTCCAGATGCATCTCTCGACATGAGTGTGTTCGACGCGATCGACACG 630

QY 601 SerGlyIleSerMetLeuGluGluValLysIleIleThrGluArgGluLeuGlnLeu 620
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 629 AGTGGCTCGATGCATCTGCAGAGCTGAGAAGGTGCTGGAAAAAAGAAACATCGAGCTT 570
```

```
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db 569 GTGCTGCCAACCCGGTGTGATCGGTACCGAGAGGCTGTACAACTCGGTGGTGGCAAG 510
Qy 641 HisLeuGlyLysLysTspIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPhe 660
Db 509 ACCTTCGGCTCAGACCGGTGTCTTCAGGCTGCCGAGCGCTCGCGGCGGCG----- 456
Qy 661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGlu 673
Db 455 -----CGGCACAAACACAGCCCTGACAGAGCAG 426

RESULT 13
US-10-437-963-77012
; Sequence 77012, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77012
; LENGTH: 4390
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRI4530_76952C.1
US-10-437-963-77012

Alignment Scores:
Pred. No.: 1,29e-167 Length: 4390
Score: 1698.00 Matches: 328
Percent Similarity: 72.61% Conservative: 120
Best Local Similarity: 53.16% Mismatches: 164
Query Match: 48.86% Indels: 6
DB: Gaps: 2

US-10-762-049-18 (1-680) x US-10-437-963-77012 (1-4390)
Qy 42 HisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLys 61
Db 503 TACAATGTTCGGCGCGCGCGCGAAGAACCTCTCTGGCGGAGTTCGCGCGCACGGTGAAG 562
Qy 62 GluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLys 81
Db 563 GAGACGTTCTTCTCCGACGAGCGATGCGCGGTGTACAGGACGAGCGAGGTCCAGGAAG 622
Qy 82 PheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPhe 101
Db 623 CTATGGCTCGCCTTGACGACGCTTCCCGGTGTTCGATGGGCGCAGACATACACCTC 682
Qy 102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db 683 GCCAAGTTCAGGCGCACCTCAITGGCGGCTCACCTTCCAGCCTCGTCATACCTCAG 742
Qy 122 GlyLeuSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPhe 141
Db 743 GACATCGGTACCGAAGCTTGTAACTCCACGAGATGGGTGGGTGACAGTAGCTTC 802
Qy 142 IleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db 803 GTCCCGCCATTGATATACGCTCTGTATGGGCGACCTCAAGGAGGAGTAGCAATGGGTCCAGTG 862
```

```
Qy 162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db 863 GCGGTCTATCTACTGCTGCTGCTACTCTCTCCAGGAGGAGATTGACTCAAAAGAGAAC 922
Qy 182 ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db 923 CGCTAGATTACAGACGGTTCGCTTCACAGCAGCTCTCTTTCGAGAGTTCACAGAGCG 982
Qy 202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db 983 GCGCTGGTTCCTGCAGGCTAGGTTTCATCATGCGTCTCTGCTCATGCTGCCATCATC 1042
Qy 222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db 1043 GGATTCATGGCGCGCGCCCATCACCATTGCTCTTTCAGCAGCTTAAAGCTTCCTTGA 1102
Qy 242 LeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln 261
Db 1103 ATTGCNAACTTACCAAGAAGACTGACATCTCGTTCATGAATCAGTCTCGGGNAAT 1162
Qy 262 ThrHis-----GluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheP 279
Db 1163 GTTCACCATGGGCAATGGAATGCGCAGACAATATTGATCGGAGCATCATTTTGGCAT 1222
Qy 279 eLeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMe 299
Db 1223 CCTCTCGGT-TGCCAA-----GGCAAGAAGAACAGAGCTCTTCGGGTCCCAACA 1375
Qy 299 tAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGlu 319
Db 1276 TGCACCATCATTTTGGTGATCATTTCAACTTTGCTGCTACATCATCTCGTCTGACAA 1335
Qy 319 sHisGlyValGluValIleGlyLeuLysLysGlyLeuAsnProProSerLeuThrAs 339
Db 1336 ACAAGCGTGCCAATCGTAAAAACGTCAGAAAGSCATCAATCCACCTTCAGCTAGCCT 1395
Qy 339 nLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValGlyIle 359
Db 1396 GATATTTTTCCTGGCCCATCTCTCTCAAGGATTCAAAATTTGGAGTAGTAGCTGGA 1455
Qy 359 eIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAs 379
Db 1456 GATAGCCCTTACGAAAGCGATTGCGATTGGAAGAACATTTTCTGGATTGAACGATTACCA 1515
Qy 379 nIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPhe 399
Db 1516 GATAGATGGAAACAAAGAAATGTGGCTCTAGGAACCATGAATGTGTCGTTCAATGAC 1575
Qy 399 rSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCy 419
Db 1576 GTCTTGCTATATAGCCACAGGTGGTTCGACGATCAGCAGTCAATTCATGCTGGAGG 1635
Qy 419 sLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLe 439
Db 1636 TAAACACCAATGTCCAAATTTGTTATGTCACTGTAGTATTGCTTGCATCTCTGTGGAT 1695
Qy 439 uThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMe 459
Db 1696 CACTCCATGTTTCAAGTACACGCCCAATGCCACCATTTCTTCCATCATCATATCAGCAG 1755
Qy 459 tLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPh 479
Db 1756 GCTTGGCTTATTGATTTGAAATTCAGCTTACCTTATCTCGAAGGTTGATAAGTTGAC 1815
Qy 479 eValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVa 499
Db 1816 CATGGGCTTGGGGGCAATCTCTGGAGTAATATTTTTCATCTGTGGAGTAGTGGCTTGT 1875
Qy 499 tIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPh 519
Db 1876 CATTCGGGTGTAAATATCACTAATCAAACTTCTGCTCCATGTAAACACGCCCAAGGAC 1935
Qy 519 eValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAl 539
```



```
Db 1936 TTTACTTGGCAACCTTCCAGAACGATTATCTATAGAAATGTTGAACAATATCCAGAGC 1995
Qy 539 aLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAl 559
Db 1996 TACCAAGTTCGCCAGGATGCTAATTGTAGAGTGGACTCAGCAATATATCTTCACAAACTC 2055
Qy 559 aSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAl 579
Db 2056 CAATTATGTTAAAGAAAGATGCTGAGATGGCTGAGAGATGAGGAAGACATCAAAAGGA 2115
Qy 579 aThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAs 599
Db 2116 ACAGAAGTTACCAAAAATTGAGTTTCTGATTGTTGACCTATCTCTGTTAAATGATATTGA 2175
Qy 599 pThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgGluLeuGl 619
Db 2176 CACAAGTGGAGATCCATGCCCTTCAAGAGATTGTTGAGGACACTTGAAGAGCGCCAGATTCA 2235
Qy 619 nLeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGl 639
Db 2236 GCTGATTTTCGCCAATTCCTGGGGGGCTGTGATCCAAAGCTCCGGTCAGCAAAATTCAC 2295
Qy 639 nAsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluAlaVal 655
Db 2296 AGAGCTCATTTGGTGAAGAAAGATATGCTCTGACAGTTGGTGACGCCGTA 2344
```

RESULT 14

```
US-10-424-599-131117
; Sequence 131117, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131117
; LENGTH: 1518
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89405C.1
US-10-424-599-131117
```

```
Alignment Scores:
Pred. No.: 4,22e-166 Length: 1518
Score: 1676.50 Matches: 344
Percent Similarity: 85.82% Conservative: 19
Best Local Similarity: 81.32% Mismatches: 28
Query Match: 48.24% Indels: 34
DB: 18 Gaps: 2
```

US-10-762-049-18 (1-680) x US-10-424-599-131117 (1-1518)

```
Qy 258 ValPheThrGlnThrHisGluTrpA-gTrpGluSerAlaValLeuGlyCysValPheIle 277
Db 51 GTTTTTTCCAAACGCATGAGTGGAGTGGGAAAGCGCTGTGTGGTGTGTTGCTTTAAT 110
Qy 278 PhePheIleuLeuSerThrArgTyrPheSerLysLysArgProArgPheThrValSer 297
Db 111 TTCTTCTCTCTCAGAGATACTTCAGCAAAACGACGCCAAAGTCTCTCTGGGTGTCA 170
Qy 298 AlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAla 317
Db 171 GCATGGCCACCATGACGCTCCGTTATATGGGAAGTCTGCTGCTCTATGTCACACAGCT 230
Qy 318 GluLysHisGlyValGluValIleGlyGluLeuLysGlyLeuAsnProProSerLeu 337
```

```
Db 231 GAGAAACATGGAGTTCAAGTGTATAGAAATTTGAAGAAAGGGGCTGAATCCACTTCGGCA 290
Qy 338 ThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValVal 357
Db 291 ACGGATTTTGGTGTGTGTGTACCTTACATGGGCACTGCCATAAAACCTGGATTCGTCACT 350
Qy 358 GlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsn 377
Db 351 GCATTTATAGCACTTCCGGAAGGATAGACAGTGGGAAGAGCTTTTGAATGTTTAAAAAT 410
Qy 378 TyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySer 397
Db 411 TATCATATTGATGGCAACAAAGAGATGATAGCTATTGGAAACCATGAATATTTCGATCG 470
Qy 398 PheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAla 417
Db 471 TTCCTTCTTGTCTATCTCACCACTGGACATTTTCGGGTGCTGCTATCAGCCATTATTGTATCT 530
Qy 418 GlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeu 437
Db 531 GGGTGCAGACAGCTGCATCCAAACATCATATGGCAATTGCAGTAATGTTGACATTGTTA 590
Qy 438 PheLeuThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSer 457
Db 591 TTCCTTAACACCAATTGTTCCATTACACTCCCTCGTGTGTGCTATCAGCCATTATTGTATCT 650
Qy 458 AlaMetLeuGlyLeuIleAspTyrGluAlaIleHisLeuPheLysValAspLysPhe 477
Db 651 GCAATGCTTGGCCTCATAGATTATGAGAGCCATCCATCTATATTAAGTCGCACAAATTT 710
Qy 478 AspPheValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGly 497
Db 711 GACTTTGTCGTGTCATGAGTCATACGTTGGCGGTGCTTTGGCAGTGTGAAATTTGSC 770
Qy 498 LeuValIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArg 517
Db 771 TTAGTCATAGCTATTGGCGATATCTGACTTCGGGTACTTCTATTATTTCGAAGGCGAG-- 828
Qy 518 ThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGln 537
Db 829 ACATTTGTTTTGGGAACATTCCAAATTCGTGTGATATACAGAAATGTTGAGCAGCTATCCA 888
Qy 538 AsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAla 557
Db 889 AATGCAAAACATGTTCTTGGAAATGCTAATTCATAGAGATTGATGCACCAATTTTACTTTGGC 948
Qy 558 AsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIle 577
Db 949 AATGCAAGCTATTTAAGAGAAAGGATCA--CGATGGATTGATGAAGAGAGAAAGAAATC 1006
Qy 578 LysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsn 597
Db 1007 AAAGCT-----GGTATG 1018
Qy 598 IleAspThrSerGly-IleSerMetLeuGluGluValLysLysIleThrGluArgArgI 617
Db 1019 TTTGATATGAGTGTGACGTGTATGCTTGGGAGATT----- 1055
Qy 617 uLeuGlnLeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerIy 637
Db 1056 -----CCAAAGAGCGAGGTGATGAAGAGAGCTGAACAAATCCAA 1093
Qy 637 sPheGlnAsnHisLeuGlyLysTrpIleTyrLeuThrValGluGluAlaValGlyAl 657
Db 1094 GTTCTCTGATGAACCTGGGGCAAAATGGATCTATCTGACAGTTGAAGAGGCTGTGGTGC 1153
Qy 657 aCysAsnPheAsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTr 677
Db 1154 GTGCAACTTCATGCTCCATTCCTACAAACCGAACCCCATGAAGGATGAATCAGAAGGTTG 1213
Qy 677 pAsnAsn 679
```


Qy	580	ThrGlyGluThrSerLeuGlnTyrValIleAspMetSerAlaValGlyAsnIleAsp	599
Db	1696	GTCGCCGCGAGGACCTGCTCTACGTGCTCGACATCGCGCGCGTGAACGGCGATCGAC	1755
Qy	600	ThrSerGlyIleSerMetLeuGluValIysIysIleThrGluArgArgGluLeuGln	619
		:::	
Db	1756	AGCCCCGGATCGAGTCTCGGGAGGTGCACGGCGAGCTGGAGAGGAAGGGATGAAG	1815
Qy	620	LeuValLeuValAenProValSerGluValMetIysLysLeuAsnLysSerLysPheGln	639
		::: ::	
Db	1816	ATGCGCGTGAACACCCGAGATCGCGTGGCGGAGAAAGCTGCTGTCTCGCGGCTCGCG	1875
Qy	640	AsnHisLeuGlyLysIysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn	659
		:::	
Db	1876	GAGCTGTCGGCGAGACTGGATGTTCTCTCAATGGCGACGCTTGGCCGCGTCCCGG	1935
Qy	660	PheAsnLeuArgAlaSerLys	666
		:::	
Db	1936	TACACGCTCCAGGCGCTCCAAG	1956

Search completed: August 31, 2005, 16:25:53
Job time : 1182 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 31, 2005, 11:52:07 ; Search time 5789 Seconds

(without alignments)
4471.187 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNIT.....NLRASKTNPKKDETCGNNV 680

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10762049/runat_30082005_121245_5984/app_query.fasta_1.839
-DB=EST -OPMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10762049 @CGN 1.1 3437 @runat_30082005_121245_5984 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1.*
2: gb_est2.*
3: gb_hrc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_ges1.*
9: gb_ges2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	2416	69.5	2158	3 CNS0A7F1 BX822461 Arabidops
2	1845.5	53.1	2279	3 AY105934 Zea mays
3	1813.5	52.2	2001	9 CL959635 OsIFCC003
4	1750	50.4	2221	3 CNS09YVW BX832405 Arabidops
5	1720.5	49.5	1812	9 CL969864 CL969864 OsIFCC019
6	1698	48.9	1959	9 CL961102 CL961102 OsIFCC005
7	1474	42.4	2321	3 CNS0A1L2 BX832587 Arabidops
8	1301	37.4	854	6 CB8892638 CB8892638 EST645430
9	1223.5	35.2	982	7 CK269206 CK269206 EST715284

10	1135	32.7	816	6 CB624151
11	1122	32.3	828	7 CV130479
12	1078	31.0	831	7 CV107436
13	1016	29.2	706	7 CO102483
14	1013	29.2	783	7 CK256297
15	1000	28.8	785	7 CF438329
16	994	28.6	768	7 CN190878
17	983	28.3	993	6 CA133402
18	970	27.9	583	2 AW598447
19	961.5	27.7	749	1 AJ794089
20	955.5	27.5	1534	5 BU693126
21	942	27.1	761	5 BQ851246
22	930	26.8	891	6 CD441698
23	929.5	26.7	985	9 CG279198
24	924	26.6	738	5 BU025764
25	922	26.5	682	6 CA484924
26	917	26.4	1043	7 CK163232
27	916	26.4	594	7 CO982632
28	916	26.4	673	7 CV474471
29	915	26.3	704	6 CA220343
30	911	26.2	728	5 BU025129
31	910	26.2	745	5 BQ996962
32	908	26.1	652	6 CB035946
33	904	26.0	728	2 AW036000
34	902.5	26.0	985	9 CG262551
35	900	25.9	666	7 CF438196
36	899	25.9	697	5 BQ850117
37	897	25.8	699	5 BQ849576
38	893.5	25.7	968	9 CG323044
39	882.5	25.4	778	8 CC005901
40	881	25.4	675	7 CK901981
41	876	25.2	656	5 BU101062
42	875	25.2	866	7 CF519225
43	862.5	24.8	968	6 CA195691
44	851	24.5	683	5 BQ970363
45	848	24.4	605	6 CA134297

ALIGNMENTS

RESULT 1

CNS0A7F1

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CNS0A7F1 2158 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLRFB3C02 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).

BX822461

HTC; GSLT cDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 2158)

Castell, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,

Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schacher, V.,

Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished

2 (bases 1 to 2158)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.									
FEATURES									
source	1. .2158 Location/Qualifiers /organism="Arabidopsis thaliana" /mol_type="mRNA" /strain="Col-0" /db_xref="taxon:3702" /clone="GSLTPB32C02" /tissue_type="Flowers and buds" /plasmid="pCMVSPORT_6" complement(1..2158) /gene="At3g51895"								
gene									
ORIGIN									
Alignment Scores:									
Pred. No.:	2,3e-272	Length:	2158						
Score:	2416.00	Matches:	469						
Percent Similarity:	85.02%	Conservative:	93						
Best Local Similarity:	70.95%	Mismatches:	93						
Query Match:	69.53%	Indels:	7						
DB:	3	Gaps:	4						
US-10-762-049-18 (1-680) x CNS0A7F1 (1-2158)									
Qy	24	MetGlySerValAspTyrGluTyrProLeuGlyMetAsnPheGluAtrgValHisGln	43						
Db	62	ATGGGCACGGAGGACTACATTCCTTCAGGACGGAGGAGCTTACC	CGCCGCCACCAC	121					
Qy	44	---ValGluValProProProGlnProPhePhePheLysSerLeuLysTyrSerLeuLysGlu	62						
Db	122	ACGGTGGAGGCTCCACACACCTCAACCGTCTTGAAGTCACTTCAGTACTCAGTCAAGAA	181						
Qy	63	ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe	82						
Db	182	ACTCTGTTCCAGACGACCTTTTAGACAAATTTAAGAACCAAAATGATCAAGAAAAATTT	241						
Qy	83	MetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThrPheGln	102						
Db	242	GTGTTAGGCCTCAAAATCTCTCCGATTTTGAATGGGCACACCGTACAAATCTCAAG	301						
Qy	103	PheLeuLysAlaAspLeuAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly	122						
Db	302	TTCTTCAAAATCAGATCTCATCGCGGAATCACCATCGTAGTCTCGCATCCTCCAGGC	361						
Qy	123	IleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIle	142						
Db	362	ATCAGTTACGCCAAACTTGCTAACTTGCCCCAAATCTTGCGCTTTATCGAGTTTGTG	421						
Qy	143	ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla	162						
Db	422	CCGCCATTGGTATACGGGGTGCTAGGAGTTCAAGGACTTAGCGGTGGGAAACGGTTGCG	481						
Qy	163	ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsmGlnAspPro	182						
Db	482	GTTCGCTCTCTGTTACAGGTGCAATGCTGAGCAAAAGAGTTGATGCTGAGAAAGATCCT	541						
Qy	183	LysLeuTyr-LeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAl	202						
Db	542	AAGCTTTTACCATTACCATTCGCTTTTCCACGGCCACTTTTTTCCCGCGCGCTTCTC	601						
Qy	202	a-LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleG	222						
Db	602	ATCTTGGAAATTTTCAGGTAGGGTTTCATAGTGGACTTTCTATCGCATGCAACGATAG	661						
Qy	222	lyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuSerIleLeuGlyL	242						


```
Db 1269 TGCTAGCTAAACAACAGGTGGTCTTCCCGCTCTGCTGTAACACCAACGCCGGCTCAAG 1328
Qy 421 ThrAlaAlaSerAsnIlelleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db 1329 ACTGCCATGTCACCAACGTGATCATGGCGCTGACGTGATGCTACGCTGTGTTCCCATG 1388
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIlelleValSerAlaMetLeu 460
Db 1389 CCACTGTTCTGTATACACCAACCAACGTTGCTCTGGAGCGATCATCATGCCGGGGTATC 1448
Qy 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 1449 GGCCTGATCGATTTCGCCGGGTGTACCACTCTGGAAGATGACAAGATGGATTTTCTG 1508
Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db 1509 GTGTGGCTTGGCGGTTCGCCGGGTATCTTTCATCTCAGTCCAAAGAGCCCTTGGGATA 1568
Qy 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
Db 1569 GCGTTGGTATCTATATTAGGTGTGTGATCAGATCACAAGGCCGAAGATGATGGTT 1628
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1629 CAAGGGAACATCAAGGGGACTGATATTTACAGAGACTGTCATCACTACAAGGAGGCCAA 1688
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db 1689 AGAGTTTCTGGTTCCTGATCTTGGCCATTTGAAGCAGCCGATAAATTCGCCAACTCCAAC 1748
Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db 1749 TACTGATGAAGAGATTAAAGATGGATA-----GAGGAAGAATCTTTGAAACAGGAT 1802
Qy 581 GlyGluThrSerLeuGlnTyrValIlelleAspMetSerAlaValGlyAsnIleAspThr 600
Db 1803 AAACATACTGAACTCCATTTTCATAATCTTGGATCTGTCAGTCTGCTGCAATTGACACA 1862
Qy 601 SerGlyIleSerMetLeuGluValValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db 1863 AGTGGCATAGCGTTCCTCATTTGATAAAGAAATCAATGAGAAACGTGGTCTGGAGCTT 1922
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLys---SerLysPheGln 639
Db 1923 GTGCTTGTCAATCCAACCTGAGAGATCATGAGAGAAATCAACAGTCGAACAGGCTGAA 1982
Qy 640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluAlaValGlyAlaCysAsn 659
Db 1983 AACTATTTTAGGCCAGATTGCTTGTATCTGACCACCTGGCGAAGCAATC-----GCTTCA 2036
Qy 660 PheAsnLeuArgAlaSerLysThrAsnProLysLysAsp 672
Db 2037 CTTTCTGCACTTGCCAAGATGACAAAACCCCTAAATGGAT 2075
```

RESULT 3

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

CL959635 2001 bp DNA linear GSS 21-SEP-2004
OsIFCC003265 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
CL959635
CL959635.1 GI:52373944
GSS.
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 2001)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and

its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers
1. .2001
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Alignment Scores:
Pred. No.: 1,73e-201 Length: 2001
Score: 1813.50 Matches: 341
Percent Similarity: 73.10% Conservative: 121
Best Local Similarity: 53.96% Mismatches: 167
Query Match: 52.19% Indels: 3
DB: 9 Gaps: 3
US-10-762-049-18 (1-680) x CL959635 (1-2001)

```
Qy 43 GlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 106 GAGGTGAACCTGTCCGGCGCGCCGTTCCGGGAGAGACTGTGTTCGACCTCGCGGAG 165
Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 166 ACGTTCTTCCCGAGACCCCGTTCCGGGGTTCGGCGGCTCCGCCCGCCAGCGCGCG 225
Qy 83 MetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThrPheGln 102
Db 226 TGGTCGGCGGTCAAGTACTTCTCCGGCGCTGGACTGGGTCCCGGCTACCGCTCGAC 285
Qy 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 286 AAGTTCAAGTTCGACTCTCTCCGCGCATCATCATCGCAGCTCGCCATCCCGCAGGCG 345
Qy 123 IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIle 142
Db 346 ATCAGCTACGCCCGCTCGCCAACTCCCTCCAATCATCGGCTCTATTTCGAGCTTCGTG 405
Qy 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 406 CCGCGTTGATGTACGGGTGTTCGGGAGCTCCAAACACTCGCGGTGGGACGGTGGCG 465
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
Db 466 GCGGCGTCTGCTGCTGGCGTCCATCATCGAGACGAGGTGGCGCGGACGAGAACCGG 525
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 526 CAGCTGTACCTCGACGCTCTTTCACCGCGGCTTCTTCACCGGCTCTTTCAGACCGCG 585
Qy 203 LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleLeuGly 222
Db 586 CTAGGGGCTTTAGGTAGGGCTAATAGTGGATTTCTCTCGCGGTTCGACGATCACCGGG 645
Qy 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeu 242
Db 646 TTCATGGGCGGCACGCGGATGATCATCATCTGACGAGCTTCAAGGGCTCTCTCGGGGATG 705
Qy 243 GluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThr 262
Db 706 AAGCACTTCACCACCAAGACAGACATCATCTCCGCTCTGCTCCACTTACCATTACCGG 765
```

```
QY 263 HisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSer 282
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
766 CACAGTGGAAATGGCAGAGTCAGATTCTCGCATATGCTTCTCTTGTCTTCTGCTGTCG 825
QY 283 ThrArgTyrPheSerLysArgProArgPhePheTrpValSerAlaMetAlaProLeu 302
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
826 AGTAAGCAGCTGAAAGAACTGCCAAGTTGTTCTGGGTGTCGGCCATTCGCCATTTC 885
QY 303 ThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVal 322
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
886 ATGGTTGCTGTCAGGAGCATCTTCGCTTCTCCTCGTCAAGGTGACGAGCATGGAATT 945
QY 323 GluValIleGlyGluLeuLysGlyLeuAsnProSerLeuThrAsnLeuValPhe 342
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
946 CCGATAGTCGGTACCTGGAATAAGGATCAATCCCTGTCCTATTCGGCAGCTAACGTTTC 1005
QY 343 ValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeu 362
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1006 CAGTCCAGGCACATGAAGCGCAGTGAAGCTGGTCTCTTGTGGGATCCTAGCACTG 1065
QY 363 AlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly 382
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1066 GCAGAGGAATAGCGTCGCCGCGAAGCTTGGCGATGGTCAAGAACGAGCAGATCGACGCC 1125
QY 383 AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1126 AACAAGGAGATGATCGCGTTTGGTATCATGAACATCCCGGTTCTTTTCACCTCTGCTAC 1185
QY 403 LeuThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla 422
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1186 CTCACGACAGCGCGTTCTCAAAGTCGGCGTGAACCTTCACCGCGGGTGCAGAGCCGC 1245
QY 423 AlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeu 442
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1246 ATGTCGAACGTGTGATGTCGGTGCATCATGCTGGTGTACTGTCTCGGCACCCCTC 1305
QY 443 PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu 462
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1306 TTCAAAGTACACCCCGCTGTGTGGCGTCTCTCCATCATCTGCTCGCATGATCGGCTC 1365
QY 463 IleAspTyrGluAlaIleIleLeuPheLysValAspLysPheAspPheValValCys 482
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1366 GTCAAGTCAAGAGGTTCTGCCACCTCTACAAGTCGACAAAGTTCGACTTCTCGCATGTC 1425
QY 483 MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle 502
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1426 ATGTCGCTCTCTCGCGCTGCTCTTCTTCCACCATGTTCGACGCTCAGCGCATCCGTA 1485
QY 503 ValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGly 522
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1486 GGCTTGTGCTGTCAGGGCGTGTGCTGACGTGGCCAGGCTGCCACCTCAAGCTCGGG 1545
QY 523 AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVal 542
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1546 AACATAGCAGGAGCAGACCTTCCGCGAGTGCAGCAATACCCCAAGGCAAGAGCATC 1605
QY 543 ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu 562
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1606 CTTGGCATCTCTGCTTGCAGCTCGGCTCTCCCATCTACTTCTCATGCGGGTTACCTG 1665
QY 563 ArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGlu 582
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1666 CGAGAAAGGATTTTGATGGTGAAGATGAGATAACCTGTGCAAGAGCGTGGGCGCAC 1725
QY 583 ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly 602
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1726 ---GATTGCAATCTGGTTCCTTGATCTGGTGGTGTCTACTTCTGTCGAACACTCTGGA 1782
QY 603 IleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLeu 622
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1783 GTTGGGATGCTACTAGAGGTACACAGAGGCTCGAACGAAGAGGGATCAGCATAGTCTCTG 1842
```

```
QY 623 ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1843 ACGAATCCGAGGCTAGAGGTGCAGAGAAGCTGGTCTGTCTGGATACGTCAGGGACATC 1902
QY 642 LeuGlyLysTrpIleTyrLeuThrValGluAlaValGlyAlaCysAsnPheAsn 661
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1903 TTAGGGATGAGTGGGTCTCTTGACGGTCAAGGACGCCATCAGCGGTGCTCATACGGCG 1962
QY 662 LeuArgAlaSerLysThrAsnProLysLysAspGlu 673
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1963 CTCAGATATCCAGA---AATAAGGAGGAAGACGAA 1995

RESULT 4
CNS09YVW 2221 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTPGH702F01 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX832405
VERSION BX832405.1 GI:42455111
SOURCE HTC; GSUT_cDNA.
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 2221)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2221)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?ggb?source=Arabidopsis.

FEATURES
source
Location/Qualifiers
1..2221
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH702F01"
/cisue_type="Hormone Treated Callus"
/plasmids="pCMVSPORT 6"
complement(1..2221)
/gene="At5g19600"

gene
ORIGIN
Alignment Scores: 6.12e-194 Length: 2221
Pred. No.: 1750.00 Matches: 334
Score: 72.13% Conservative: 119
Percent Similarity: 53.18% Mismatches: 158
Best Local Similarity: 50.36% Indels: 18
Query Match: 3 Gaps: 4
DB:
```

```
US-10-762-049-18 (1-680) x CNS09YVW (1-2221)
QY 44 ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThr 63
Db 127 GTRAAATTCGACCCCGAGAGCTTTGGTTCAAAATTTAAGTCAAAATGTAAGAAACC 186
QY 64 PhePheProAspAspProLeuArgGlnPheLysAsnLysPro-----AlaSer 79
Db 187 TTCTTTCCGATCATCAATCAAAACCGATATCGCAAGAGCCAAACCGTTTGGCTAAAAACA 246
QY 80 LysLysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluThrAlaProLysTyr 99
Db 247 AAGAAGTTA-----TTGGAGTATTTTGTGCCAATCTTTGAGTGGCTCCCAAGTAT 297
QY 100 ThrPheGlnPheLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIle 119
Db 298 GATATGCAAAACTAAAGTATGATGCTCGCCGGAATCACCATCACCAGTCTTGCTGT 357
QY 120 ProGlnGlyLeuSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSer 139
Db 358 CCTCAGGCATCAGTATGCTAAGTCGCGAGTATCCCTCCCATCATAGGCTATATTCA 417
QY 140 SerPheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGly 159
Db 418 AGTTTGTGCGGCTTTTGTGTATGCGTGTTCGGAAGTTTCGAACAATCTGGCAGTGGGA 477
QY 160 ThrValAlaValGlySerLeuMetGlySerMetLeuSerAsnAlaValAspProAsn 179
Db 478 ACAGTCGCGAGCATGCTTTTGTGATCGTCGAGACATTCGGTGAGAGATGATAAAAAAT 537
QY 180 GluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPhe 199
Db 538 GAG--CCTGAACCTCTACCTTCATTTGATCTTCACCGCCACTCTCATCACCGGTCTATT 594
QY 200 GlnAlaAlaLeuGlyLeuPheArgLeuGlyLeuValLeuValAspPheLeuSerHisAlaThr 219
Db 595 CRAATTTGTATGGGATTTCTGAGTTGGGATATGTTGGTATTTCTGTGCACACTCGACC 654
QY 220 IleIleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIle 239
Db 655 ATACCGGCTTCATGGCGGTACCGCAATCATCTCTCTACACAGCTTTAAGGCAATC 714
QY 240 LeuGlyLeuGlnHisPheThrHisGlyAlaAspIleSerValMetArgSerValPhe 259
Db 715 TTTGGACTAGTCCATTTTACGCACAAACCGATGTCGTATCAGTTCTCCACTCAATCTTA 774
QY 260 ThrGlnThrHisGluThrArgThrGluSerAlaValLeuGlyCysValPheIlePhePhe 279
Db 775 GACAAACCGGCTGAGTGGAGTGCGAAAGTACCTTGGCTGGTCTCTGTCTTTCTGTATT 834
QY 280 LeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheThrValSerAlaLeu 299
Db 835 CTACAATCCACTCGATACATAAGCAAAAGTACCCAAAGCTATTTTGGGGTATCAGCAAT 894
QY 299 tAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLys 319
Db 895 GGGTCCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 954
QY 319 sHisGlyValGluValIleGlyLysLysGlyLeuAsnProProSerLeuThrAs 339
Db 955 ACACGGCATTCCTACCTGTTGGGCTTTAAAGAAAGGACTAAATCCCTCTTATTCACTT 1014
QY 339 nLeuValPheValSerProTyrMetThrAlaValLysThrGlyIleValValGlyLys 359
Db 1015 -TTGAATTTGATCCCAAGTATCTAGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1073
QY 359 eIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAs 379
Db 1074 CATCGCTTTGGCGGAAGGATAGCAATCGAAGAACCTTTGCTGTAATGAAGAACGAACA 1133
QY 379 nIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheTh 399
```

```
Db 1134 AACGACCGGACAAAGAGAGATGATGACATTTGGTCTTATGAACGTAATGGCTCTCTAC 1193
QY 399 rSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCys 419
Db 1194 TTCCTTGCTATTTGACACCGGGCCATTTTCAAGACAGCAGTGAACACTACACGCGGGAC 1253
QY 419 sLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeu 439
Db 1254 AAAGACACCGATGTCGAATGTAAGTAAAGCGGCTTTGGCATGATGCTTGTCTCTCTCT 1313
QY 439 uThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaLeu 459
Db 1314 ACGGCTCTATTACGCTACACACCATTAAGTGGTCTCTCAGCCATCATTTATGACCCAT 1373
QY 459 tLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPh 479
Db 1374 GTTGGGTCTCATCACTACGAGGAGATGTACCATCTCTTCAAGGTTGACAAATTTGATT 1433
QY 479 eValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVa 499
Db 1434 CCTCGTCTGATGTCGCTTTTTCGGTGTTCATTTCTTAGCATGGACTACGGTCTCAT 1493
QY 499 lIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPh 519
Db 1494 CATCTCGGTGGATTTCTATCGTAAGAGCTTTGTGTACGTTGCAAGGCCCTCCACTTG 1553
QY 519 eValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAl 539
Db 1554 TAAATTTGGAGGATACCAACTCGTGTATGTTTAGACATAGAACAGTACCTTGCCTC 1613
QY 539 aLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAl 559
Db 1614 CGAGGAGATGCTGCTGTATATCATTTCTCAATAGGCTCTCCAGTCTTTTTCGCAATAG 1673
QY 559 aSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluGluArgIleLysAl 579
Db 1674 CACTTACGTCGCGTAACCGATCTTCAAGTGGATCGTGAACCTGAA----- 1722
QY 579 aThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAs 599
Db 1723 -----GCCATCGAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1769
QY 599 pThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlu 619
Db 1770 CATGACCGGATGGAACACTACTCGAAATTCAGAGAATCTCTGGATCAAAAACATCAA 1829
QY 619 nLeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGl 639
Db 1830 GATGTGATATAAATCCAAAGATTTGAAAGTCTTGAAAGAGTGTGTTGTCCTCATTTCT 1889
QY 639 nAsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAs 659
Db 1890 GGAGAGATAGGAAAGAGTATATGTTTATCGATGACGATGACGATGACGATGACGATG 1949
QY 659 nPheAsnLeuArgAlaSerLys 666
Db 1950 ATTTAATCTTACCACCACCAAG 1971

RESULT 5
CL969864 1812 bp DNA linear GSS 21-SEP-2004
LOCUS OsIFCC019054 Oryza sativa Express Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL969864.1 GI:52394347
VERSION GSS.
KEYWORDS Oryza sativa (indica cultivar-group)
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 1812)
AUTHORS Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
```


Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis

JOURNAL COMMENT

Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES

source
1. .1812
/organisms="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/notes="Oryza sativa exon trapped genomic sequences"

ORIGIN

Alignment Scores:

Pred. No.: 1,31e-190 Length: 1812
Score: 1720.50 Matches: 335
Percent Similarity: 71.20% Conservative: 105
Best Local Similarity: 54.21% Mismatches: 149
Query Match: 49.51% Indels: 29
DB: 9 Gaps: 3

US-10-762-049-18 (1-680) x CL969864 (1-1812)

```

QY 41 ValHieGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
DB 10 CTACACAAGGTGGCGCGCGCGCGCAGCAGCAGCGCAGCTGAAGCGGAGGTG 69
QY 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
DB 70 AAGGAGACCTTCTTCCCGCAGCAGCGCGTTCGCGGGTTCAAGGGGAAGCGCTGACGACC 129
QY 81 LysPheMetLeuGlyLeuGlnPhePheProLeuPheGluThrAlaProLysTyrThr 100
DB 130 AAGTGGTCTAGCGGTGGTACGTACCTGTTCCCGATCTCCGACGTG----- 174
QY 101 PheGlnPheLysAlaAspLeuAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
DB 174 ----- 174
QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProPheLeuGlyLeuTyrSerSer 140
DB 175 ---GGCATTAGCTACGCGAAGCTGGCGAGCTTCCCTCCGATCATCGCGCTCTATTGAGC 231
QY 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
DB 232 TTGTGCGCGCGAGTGGTACGCGGTGCTGGGAGCTCGAGGACCTGGCGGGGGCGCG 291
QY 161 ValAlaValGlySerLeuMetGlySerMetLeuSerAsnAlaValaAspProAsnGlu 180
DB 292 GTCTCGATCGCTCGTGAATAATGGGTTCATGCTGCGGCGAGCGGTGAGCCCGCGCC 351
QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
DB 352 GAGCGCGTGTCTTCTCCAGCTCGCGCTTCCACCTCCACCTTCTTCGCGCGCGCTCGTCCAG 411
QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
DB 412 GCCTCCCTCGGCATCTCCGCGTGGGGTTCATCATGATTCTCTCGAAGCGCGAGCTT 471
QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
DB 472 GTCGGGTTCATGCGCGCGCGCGATCATCTGTCGTGCGAGGAGCTCAAGGGCGCTGCTC 531

```

```

QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
DB 532 GGCATCGTCCACTTCCACCCGAGATGGCGCTCGTCCCGCTCATGCCCTCCGTCATCCAC 591
QY 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280
DB 592 CACACCAAGAGTGGTGGCGAGCATCTGATGGCGGTGCTTCTCGTGGTCTCTG 651
QY 281 LeuSerThrArgTyrPheSerLysAspProArgPhePheTrpValSerAlaMetAla 300
DB 652 CTCACGGCGAGGCATGTGATGATAAATGGCCAAAGCTTCTTGGGTTTCAGCTTGTGCT 711
QY 301 ProLeuThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHis 320
DB 712 CCTTGGCATGTGTTATCTCTACCTGCTGTTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 771
QY 321 GlyValGluValIleGlyLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
DB 772 GGCATTAGCATTAATGGCGAGCTCAATATGGCTGAATGGCGCTTCTGTTCTGTTCTGTTCT 831
QY 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
DB 832 CTCTTTGATCCTCAATATTTGGCGCTCACTGTAAGACTGGCCTTGTCAACGGAATCATC 891
QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
DB 892 TCCCTAACGGAAGGAGTGGCAGTTGGTAGGACATTTTCTTCACTCAAGGACTACCAAGTA 951
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
DB 952 GATGTGAACAAGAGATGATGGCCATAGGATGATGAACATTTGGGGCTCATGTATCATCA 1011
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
DB 1012 TGTATGTAACATACAGGACATCTCTCTGTTCTGTTGTAACACCAATGCTGTTGTTCAAG 1071
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
DB 1072 ACTGCCATGTCAATATGATCATGTCAGTGTGCTGTCACACTGTGTTGTTGTTGTTGTTG 1131
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
DB 1132 CACTGTTTGTGTACACACGATGTTGTTCTTGGAGCAATCATATTTGCTGCTGTTGATT 1191
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
DB 1192 GGGCTCATTTGATCTCCCTGCTGTATACAACATCTGGAAGATGGAAGATGGATTTCTT 1251
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
DB 1252 GTGTGTCTATGTGCAATTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1311
QY 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
DB 1312 GCGTTTGTGTATCATATTTTCAAGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1371
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
DB 1372 CAAGGAACATCAAGGACCTGATATTTACCGAAATCTTTCATCAGTACAGGACGCTCAA 1431
QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
DB 1432 AGAGTTTCTTGGGTTCTTAATCTCGAGTTGAGGCTTCTTAAACTTTTTCACCAACACCA 1491
QY 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
DB 1492 TACCTTAATGAAGGANTTAAAGATGGATA-----GAAGAAGAAAGTTCTCGCAGGAACC 1545
QY 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
DB 1546 AAACAAGAGTAAGTCCATTTCTGTAATTTTGGATCTGTCAGCTGCTCCTCAATTGATACA 1605
QY 601 SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuLeu 620

```

Db	1606																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												</
----	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Qy 479 eValValCysMetSerHlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVa 499
 Db 1395 CATGGCGTGTGGGGCAATTCCTGGAGTAATATTTCATCTGTGGAGTATGGCTTGCT 1454
 Qy 499 lIleAlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPh 519
 Db 1455 CATGGCGTGTATATCATCAATCAAGATTCGTCTCCATGTACAGCGCCAGGACAGC 1514
 Qy 519 eValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAl 539
 Db 1515 TTTACTTTGGCAACCTTCCAAAGAAGATTATCTATAGAAATGTTGAACAATATCCAGAAGC 1574
 Qy 539 aLysHisValProGlyMetLeuIleGlyValIleAspAlaProIleTyrPheAlaAsnAl 559
 Db 1575 TACCAAGGTCCAGGGATGTAATGTGAAGAGTGGACTACGCAATATACTTCAAACTC 1634
 Qy 559 aSerTyrLeuArgGluArgIleThrArgTyrPileAspGluGluGluArgIleLysAl 579
 Db 1635 CAATTATGTTAAAGAAAGATGCTGAGATGCTGAGAGATGAGAGAGAACATCAAAAGGA 1694
 Qy 579 aThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAs 599
 Db 1695 ACAGAAGTTACCAAAATTCAGTTTCTGATTGTTGACCTATCTCTGTTAAATGATATGA 1754
 Qy 599 pThrSerGlyIleSerMetLeuGluGluValLysIleThrGluArgGluLeuGl 619
 Db 1755 CACAAGTGAATCCATCCCTTCAAGAGATGTTGTGAGGACACTTGAAGAGCGCCAGATTCA 1814
 Qy 619 nLeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGl 639
 Db 1815 GCTGATTTCCCAATCTCGGGCGGCTGTGATCCAAAGACTCCGGTCAGCAAAATTCAC 1874
 Qy 639 nAsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaVal 655
 Db 1875 AGAGCTCATTTGGTGAAGAAAGATATGCTGACAGTTGGTGACGCCGTA 1923
 RESULT 7
 CENS0A1LZ
 LOCUS
 DEFINITION
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 GSLTPQH22E06 of Hormone Treated Callus of strain col-0 of
 Arabidopsis thaliana (thale cress).
 BX832587
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 2321)
 Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 Unpublished
 2 (bases 1 to 2321)
 Genoscope.
 Direct Submission
 TITLE
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 The sequences are based on single pass reads.
 Life technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banquet_BF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
 FEATURES
 source
 1..2321
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTPQH22E06"
 /tissue_type="Hormone Treated Callus"
 /plasmid="pCWSPORT_6"
 1..2321
 /gene="At5g10180"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,04e-161 Length: 2321
 Score: 1474.00 Matches: 300
 Percent Similarity: 68.47% Conservative: 130
 Best Local Similarity: 47.77% Mismatches: 185
 Query Match: 42.42% Indels: 14
 Db: 3 Gaps: 8
 US-10-762-049-18 (1-680) x CENS0A1LZ (1-2321)
 Qy 48 ProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThrPhePheProAsp 67
 Db 246 CACCTAGCCCGTGGCATGAG---CTCAAAAGACAGTCAAGGCTCTTTCCTA----- 236
 Qy 68 AspProLeuArgGlnPheLysAsn-----LysProAlaSerLysLysPheMetLeu 84
 Db 297 ACCAAAGCCAAAGAGTTCAAGTCACCTTCAAAACAGCCCTTCCAAACAAATCTCTCT 356
 Qy 85 GlyLeuGlnPhePhePheProIlePheGluTyrAlaProLysTyrThrPheGlnPheLeu 104
 Db 357 GTCCTCCAGCCATTTTCCCAATCTTCGGTGGCAGAAACTATATAACTCACCATGTTTC 416
 Qy 105 LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer 124
 Db 417 AGAAGCATCTCATGGCTGTTTAAACCTCGTAGCTCTGCATTCGCGAGCAGCATTTGGT 476
 Qy 125 TyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIleProPro 144
 Db 477 TATGCAACTCTTGCAAGGCTTGATCCTCAATATGGCTATATACGATGTGGTACCACCA 536
 Qy 145 LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly 164
 Db 537 TTGATATATGCAATTGATGGGACATCAAGAGAGATAGCAATCGGACCGGTGGCTAGTA 596
 Qy 165 SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeu 184
 Db 597 TCTTCTTATATCTTCAATGTTGAGAAACTCATCGATCCAGAAACAGATCCCTTGGGA 656
 Qy 185 TyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuGly 204
 Db 657 TACAAGAAACTGTCTCTAACCAACACTTCTTCGCGGGATCTTCCAAGCTCTTTCGGT 716
 Qy 205 LeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMet 224
 Db 717 TTATTTCAGTTTAGGGTTCTTGGTGGATTTCTGTCTGACGACGATAGTTGGGTTCATG 776
 Qy 225 GlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis 244
 Db 777 GGTGGTGCAGCCATTGTAATGGACTCCCAACAGCTTAAAGGTTGTTGGTATCATTAC 836
 Qy 245 PheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln---ThHis 263
 Db 837 TTCACCACCAACACATGACATGCTCTCTGTTCTTCGAGCTGTCTGGAGCTCTTGTCAACA 896
 Qy 264 GluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThr 283

Db	897	CAATGGAGCCCTCACACTTTTCATCCTCGGATGTTCTTCTCAGTTTATATCTTATTTACT	956
Qy	284	ArgTyrPheSerLysLysArgProArgPhePheTirpValSerAlaMetAlaProLeuThr	303
Db	957	CGCTTCATCGGAGAGAGTA-AAGAAAGCTGTTTTGGCTACCGGCAATAGCTCCGTTGATC	1015
Qy	304	SerValleLeuLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGlu	323
Db	1016	GCCTGTGTAGTGTCAACACATAATGGTGTCTGACTAAAGCCGACGAGCATGGTGTGAAG	1075
Qy	324	ValleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheVal	343
Db	1076	ACAGTGAGGCACATCAAGAGAGGCTTAATCCAAATGTCATTCAGAGATCTCGACTTAAT	1135
Qy	344	SerProTyrMetThrThrAlaValLysThrGlyLysValValGlyLysLysSerLeuAla	363
Db	1136	ACTCTCATCTCGGACAAATCGCTAAATCGGATTAATCATTTGCCATTTGTGCTCTAAC	1195
Qy	364	GluGlyLeuAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnLysAspGlyAsn	383
Db	1196	GAGCGATTTCGGTGGGAGTCTTCGCCGAATAAAGGATACACACTCGATGGAAC	1255
Qy	384	LysGluMetLeuAlaLysGlyThrMetAsnValValGlySerPheThrSerCysTyrLeu	403
Db	1256	AAAGAAATGGTGGCAATGGATTTATGAATGTTCTCGTTCCTTTCACATCTTGTACGCT	1315
Qy	404	ThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAla	423
Db	1316	GCTACTGTTTCATCTCTCGACGGCGTGAATTTTCGGCAGATGTGAGACAGCAATG	1375
Qy	424	SerAsnLeuLeuMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPhe	443
Db	1376	TCCAACATTGTTATGGCGTTACGGTGTGTAGCACTCGAGTGTCTAACGAGGCTTCTC	1435
Qy	444	HisTyrThrProLeuValValLeuSerAlaLysLysValSerAlaMetLeuGlyLeuLeu	463
Db	1436	TACTATACTCCAATCGCCATCTCGCTTCAATAATTTCTCAGCACATTCGGGCAATATC	1495
Qy	464	AspTyrGluAlaLysLysPheLysValAspLysPheAspPheValValCysMet	483
Db	1496	ACATTAACGAGGCTATTCATTTGGAAGTGTGAATAATTCGATTTCTTGCTCTCAT	1555
Qy	484	SerAlaTyrLeuGlyValValPheGlySerValGluLeuGlyLeuValLysAlaLysVal	503
Db	1556	GGAGCTTCTTTGCTGTTGTTGCTTCGTTGAGATCGACTCTCTGTGCGGGTGGTT	1615
Qy	504	IleSerValLeuArgValLeuLeuPheLysAlaArgProArgThrPheValLeuGlyAsn	523
Db	1616	ATTTCGTTTCCCAAGATCATACTCATATCAATTCGTCAGGGATAGAAACGCTTGAAGA	1675
Qy	524	IleProAsnSerValLysTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValPro	543
Db	1676	ATGCCCGGGACCGATCTTTACAGATCAATCAATATCTATGACGGTTAAGACTCCC	1735
Qy	544	GlyMetLeuLeuLeuGluLys---AspAlaProIleTyrPheAlaAsnAlaSerTyrLeu	562
Db	1736	GGAGTGTGATTTTCGTGTCAGCTCTGCATTTGTTGCTTTGCCAATGCCAGTTCAATT	1795
Qy	563	ArgGluArgIleThrArgTyrIleAsp-----GluGluGluGluArgLysAlaThr	580
Db	1796	GAGAAAGGATTATGGGATGGGTGATGAGGAAGAAGAAGAAAAACAAAGAGCAAT	1855
Qy	581	GlyGluThrSerLeuGlnTyrValLysLysAspMetSerAlaValGlyAsnLysAspThr	600
Db	1856	GCCAAAGAGAAGATCCCTCTTTGTAGTCTCTGATATGTCAGATTTGATCAACGTCGATACA	1915
Qy	601	SerGlyLysSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeu	620
Db	1916	TCGGGGATTTACTGCTTGTGGAGATTCGATAACAAATAATCAAACTGGTGTGACCTA	1975
Qy	621	ValLeuValAsnProValSer-GluValMetLysLysLeuAsnLysSerLysPheGlnAs	640

[illegible]

Db	3	ATATTGAAATGGGACCAAGCTACACTTTTCACTTTTGAATCTGAICTCATAGCTGGC	62
Qy	112	IleThrIleAlaSerLeuAlaIleProGlnGlySerTyrAlaLysLeuAlaAsnLeu	131
Db	63	ATCATATTGCTAGCTGGCTATTCTCTCAAGGCAATTAGTATGCCAACTTGGCAACCTC	122
Qy	132	ProProIleLeuGlyLeuTyrSerSerPheIleProProLeuIleTyrAlaMetMetGly	151
Db	123	CCTCCAAATCTTGACATATATTCAAGCTTTTATACCAACCGTTAATTTATGCAATGATGGGT	182
Qy	152	SerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMet	171
Db	183	AGTTCAGGAGATTGGCAGTTGGAACCTGGCAGTTGGATCTCTTCTGATGGGTTCCATG	242
Qy	172	LeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAlaPheThr	191
Db	243	TTGGCCAAATGAGGTAAATCTACCCAAATCCAAAGCTTTTCTGCACCTTGTCTTCACA	302
Qy	192	AlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGlyLeuIle	211
Db	303	GCTACATTTCTTCTGACTTTTGGCAAGCTTCATTTGGGTCTCTTTAGGTTAGGGTTATC	362
Qy	212	ValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValVal	231
Db	363	GTGGATTTCTGTACATGCAGCAATAGTAGGGTTATGGAGGAGGAGCCAGCGTGGTT	422
Qy	232	CysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAspIle	251
Db	423	TGCTACAGCAGCTGAATCAATCTTAGGACTTGACATTTACCCCATGCTGCTGATATT	482
Qy	252	IleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGluSerAlaVal	271
Db	483	GTATCGCTCATGGGCTCCGTTTATCCCAACTCATCAGTGGAGGTGGGAAAGCGTGTG	542
Qy	272	LeuGlyCysValPheIlePheLeuLeuSerThrArgTyrPheSerLysLysArgPro	291
Db	543	TTGGGATTTTGCTTCATTTCTTCTCTGTGTACAGATACATTCAGCAAAAACACCC	602
Qy	292	ArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeu	311
Db	603	AAGTCTCTTTGGGTGTCAGCAATGACGCCGTAGCGCTCGCTTATCTAGTGGCAGTTTCTG	662
Qy	312	ValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGly	331
Db	663	GTTTATTTCATCTATCGGAGCACCAGCGGGTTCAAGTGTATAGGAGAAATTTGAAGAAAGGA	722
Qy	332	LeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaVal	351
Db	723	CTAAATCCACCACTCATCAGATTTGGTATTTGTGCACCATATATGACCACAGCTATC	782
Qy	352	LysThrGlyTyrValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSer	371
Db	783	AAAACAGGCTTATTTGTCATATATAGCTCTTCGGGAAGGAAATCGCTGTGGGAAGAGC	842
Qy	372	PheAlaMetTyr	375
Db	843	TTTGCAATGTAT	854
RESULT	9		
CK269206			
LOCUS			
DEFINITION			
EST715284	potato	abiotic stress	cdna library Solanum tuberosum
clone	POACL46	5', end,	mRNA sequence.
CK269206			
CK269206.1	GI:39826184		
EST.			
ORGANISM			
Solanum tuberosum	(potato)		
Solanum tuberosum			
Eukaryota;	Viridiplantae;	Streptophyta;	Embryophyta; Tracheophyta;
Spermatophyta;	Magnoliophyta;	eudicotyledons;	core eudicots;
asterids;	lamids;	Solanales;	Solanaceae; Solanum.
1	(bases 1 to 982)		
REFERENCE			

AUTHORS
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: EST715285
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ART TAG GTG ACA CTA TAG.

FEATURES
source
1..982
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POACL46"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/note="Vector: pCMVSPORT6.1; Site 1: ECORI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cdna library. RNA sample."

ORIGIN
Alignment Scores:
Pred. No.: 1 59e-132 Length: 982
Score: 1223.50 Matches: 234
Percent Similarity: 85.08% Conservative: 34
Best Local Similarity: 74.29% Mismatches: 41
Query Match: 35.21% Indels: 6
DB: 7 Gaps: 2
US-10-762-049-18 (1-680) x CK269206 (1-982)
Qy 24 MetGlySerValAsp---TyrGluTyrProLeuGlyMetAsnAsn-----Phe 38
Db 33 ATGGGTAAACGACAGATAATTATGATACCATCAATGATGAATAATAATATGGAGAAAA 92
Qy 39 GluArg-ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTy 58
Db 93 CACAGGATACACAGAGTTGAAATTTCTCCACACACACGCTTTTTCAGTCATTGAAAAA 152
Qy 58 rSerLeuLysGluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAl 78
Db 153 TACAGTGAAGGAAACATATTATTTCCAGATGATCCATTAGCAATTCAGAACCAAAACC 212
Qy 78 aSerLysLysPheMetGlyGlnPhePheProIlePheGluTrpAlaProLy 98
Db 213 TTTGAGAAATTTATACTTTGGAGTTCAATATTTTTCATATTTTGAATGGGTTCAAG 272
Qy 98 sTyrThrPheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAl 118


```
Db      620 AACATTGGGATCACTCACTCTTGTACCTCACCACCGGTCGGTTCGAGGTCGGCC 679
Qy      413 ValAsnTyrAsnIaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaVal 432
Db      680 GTCAACTACAAACGCGGCTGCAAGACGCGGATGTCGAACGTGATCATGTCGGTGGGGTG 739
Qy      433 MetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSer 452
Db      740 ATGATCAGCTGCTGTTCTTGACGCGCTGTTCCACTACACGCGCTGGTGGTGTGCTGTCG 799
Qy      453 AlaIleIleValSer 457
Db      800 GCGATCATCATGTCG 814

RESULT 11
LOCUS   CV130479
DEFINITION X9p02a03 Populus stem seasonal library Populus deltoides cDNA, mRNA
ACCESSION CV130479
VERSION    CV130479.1 GI:51874399
KEYWORDS  EST.
SOURCE    Populus deltoides
ORGANISM  Populus deltoides
REFERENCE 1 (bases 1 to 828)
AUTHORS   Park,S. and Han,K.-H.
TITLE     Gene expression profile during seasonal growth cycle in poplar tree
JOURNAL   Unpublished (2003)
COMMENT   Contact: Kyung-Hwan Han
          Department of Forestry
          Michigan State University
          126 Natural Resources, East Lansing, MI 48824-1222, USA
          Tel: 517 353 4751
          Fax: 517 432 1143
          Email: hanky@msu.edu.

FEATURES
source   Location/Qualifiers
1..828
/organism="Populus deltoides"
/mol_type="mRNA"
/strain="ILL-129"
/db_xref="taxon:3696"
/tissue_type="stem"
/dev_stage="1 year old"
/clone_lib="Populus stem seasonal library"

ORIGIN
Alignment Scores:
Pred. No.: 1.08e-120 Length: 828
Score: 1122.00 Matches: 217
Percent Similarity: 91.18% Conservative: 31
Best Local Similarity: 79.78% Mismatches: 22
Query Match: 32.29% Indels: 2
DB: 7 Gaps: 1

US-10-762-049-18 (1-680) x CV130479 (1-828)
Qy      353 ThrGlyIleValValGly-----IleIleSerLeuAlaGluGlyIleAlaValAlaValArg 370
Db      11 ACCGGACATATGCCCGGAATTTCGGCCATTACGGCGGGGAAGGATAGCAGTAGAGAGA 70
Qy      371 SerPheAlaMetTyrIlysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGly 390
Db      71 AGTTTTCGCAATGTTCCAGATTACCATATAGATAGGCAACAAAGAGATGATTGCAATTGGG 130
Qy      391 ThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArg 410
Db      131 ACCATGAACATGTTAGAGCTCTTGACCTCTTGCTATCTCAACAGGCGCATTCGCGCA 190
Qy      411 SerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeu 430
```

```
Db      191 TCGGCTGTGAATTTCAATGCAGGATGCAAGACAGCAGTATCGAATATTGTTCATGGCGTTG 250
Qy      431 AlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValVal 450
Db      251 GCAGTCATGGTCACGTGTGTTATTTCCTTAAGCCCAATTGTTCCATTACACTCCACTTGTGGTT 310
Qy      451 LeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHis 470
Db      311 CTATCTCTCTATTATTATCTCTGCCATGCTAGGTCCTTATAGATTATGAAGCAGCAATCCAT 370
Qy      471 LeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValVal 490
Db      371 CTTTGGACTGTCGACAAGTTGACTTCATCGTGTGCATAAGCGCATATGCTGGTGTGGTT 430
Qy      491 PheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeu 510
Db      431 TTTTCAGTGTTCGAGATTGGCTTAGTCATCGCGTTCGGCAATTTCTTTGTGTCAGATTACTT 490
Qy      511 LeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyr 530
Db      491 CTCCTTCGCAAGACCAAAACATTATTCTTTGGAAACATTCCAAATTCGATGATCTAC 550
Qy      531 ArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIle 550
Db      551 AGAAATGTTGAACAGTATACGAACAACAGCAGCGTTCCCGGTGTTCTCATACTCGAGATT 610
Qy      551 AspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrrPle 570
Db      611 GATGCTCTCTACTTCTTCAATGCAAGCTACTTAAGAGAAAGGATTGCGAGGTGGTT 670
Qy      571 AspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIle 590
Db      671 GATGAAGAGAGAGAACAGTTAAATCTTCAGAGAAACCAAGCTCGAGTATGTTACTA 730
Qy      591 AspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLys 610
Db      731 GACATGGGAGCGGTAGTAACTTGCACACAGTGGGATTGGCATGCTTCGAGGAAAGTAAG 790
Qy      611 LysIleThrGluArgArgGluLeuGlnLeuValLeu 622
Db      791 AAAGTAATGGACAGAGGGAACCTCAAGCTTGTCTTG 826

RESULT 12
LOCUS   CO107436
DEFINITION GR_Eb0038K08.r GR_Eb Gossypium raimondii cDNA clone GR_Eb0038K08
3', mRNA sequence.
ACCESSION CO107436
VERSION    CO107436.1 GI:48806122
KEYWORDS  EST.
SOURCE    Gossypium raimondii
ORGANISM  Gossypium raimondii
REFERENCE 1 (bases 1 to 831)
AUTHORS   Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
          Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
          Wing,R.A.
TITLE     Global assembly of Cotton ESTs
JOURNAL   Unpublished (2004)
COMMENT   Contact: Rod A. Wing
          Arizona Genomics Institute
          The University of Arizona
          Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
          Tel: 520 626 9595
          Fax: 520 621 1259
          Email: http://genome.arizona.edu
          Plate: 0038 row: K column: 08.
          Location/Qualifiers
1..831
/organism="Gossypium raimondii"
/mol_type="mRNA"
```

```
/db xref="taxon:29730"
/clone="GR_Eb0038K08"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:
Pred. No.: 1.64e-115 Length: 831
Score: 1078.00 Matches: 201
Percent Similarity: 88.45% Conservative: 44
Best Local Similarity: 72.56% Mismatches: 31
Query Match: 31.02% Indels: 1
DB: 7 Gaps: 0

US-10-762-049-18 (1-680) x CO107436 (1-831)

QY 312 valTyrPheThrHisAlaGluLysHisGly-ValGluValIleGlyGluLeuLysGly 331
Db 2 GTTTATCTGACGATCGAGAAACATGGCCCTTCAAGTGATTGGGCACCTTGAAGAAAG 61

QY 331 YLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThrAlaVa 351
Db 62 GTTGATCCACCGTCTGCATCAGCCTGGTGTGGTGCGCCCACTTGATGGCGGCTAT 121

QY 351 llysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSe 371
Db 122 AAAACCGGATCGTCAATGGCATCATTCGCTTGTGAAGGAATTCGAGTCGGAAGAAC 181

QY 371 rPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyTh 391
Db 182 CTTTGCCATGTTCAAGAACTATCACATTCATGATGGGAAACAAAGAAATGATCGCCTCGGAT 241

QY 391 rMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSe 411
Db 242 GATGAACATTCGTGGCTCATGCACCTCTCTGTACTTAACTGCAGGACCAATTCAGAAGC 301

QY 411 rAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAl 431
Db 302 TGCAGTGAACATTTAATGCTTGAAGCAGAGAGTGCAGTCTCCAACTGTTAATGCAACAGC 361

QY 431 aValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValVal 451
Db 362 AGTGATGTTTACGTTGTATTTCTAACACCATTTGTTCCATTACACTCTCTTGTGGTACT 421

QY 451 uSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLe 471
Db 422 TTCCTCTATATATAGACGCCATGTTGGCTGCTTGACTATGAAGCTGCAATCCACCT 481

QY 471 uPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPh 491
Db 482 TTGGAATAATGACAAATTCGACTCTTCATCTGCTTGGGGGCATATCTCGGTGTGTCTCT 541

QY 491 eGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLe 511
Db 542 CGGAAGCGTTGAGATTGGTCTAATTCATTTTCGATCATCTGTTTCATTGCTAAGGATTATTCT 601

QY 511 uPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrAr 531
Db 602 CTTTGGCGGCGGCCAGACAGATGTTTTAGGCATATATTCGAACCTCTGGGATCTATAG 661

QY 531 gAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAs 551
Db 662 AAGTAGGACCAATACCAATCGCAACAGTGTTCAGGAGTTCCTCATACTGCAATAGA 721

QY 551 pAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAs 571
Db 722 CGCCCCCTGTTTCTTTTGCCAAATGCAAGCTATCTTAAGAGAAGAGATCTTCGAGATGGATT 781
```

```
571 pGluGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyr 587
Db 782 TGAAGAGGAAACACAGGTTAAATCTGCAGGAGAGCAAGCTTACACTAT 830

RESULT 13
LOCUS CO102483
DEFINITION GR_Eb0029P05.r GR_Eb Gossypium raimondii cDNA clone GR_Eb0029P05
3', mRNA sequence.
ACCESSION CO102483
VERSION CO102483.1 GI:48801169
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 706)
AUTHORS Kim H., Yu Y., Kudrna D., Hatfield J., Stum D., Mueller C.,
Udall J.A., Rapp R.A., Wendel J.F., Rao K., Soderlund C. and
Wing R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 0029 row: P column: 05.
FEATURES
Location/Qualifiers
source
1..706
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb0029P05"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."
```

```
Alignment Scores:
Pred. No.: 2.52e-108 Length: 706
Score: 1016.00 Matches: 192
Percent Similarity: 92.31% Conservative: 24
Best Local Similarity: 82.03% Mismatches: 18
Query Match: 29.24% Indels: 0
DB: 7 Gaps: 0

US-10-762-049-18 (1-680) x CO102483 (1-706)

QY 303 ThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVal 322
Db 3 ACAACGGNATTTCTTGGAAAGTCTACTCGTTTACTTGACCCCATGCTGAAAAACACACGGTGT 62

QY 323 GluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPhe 342
Db 63 GATGATTTGGAACCTTGAAGAAAGGTTGAATCCACCTCTTTTGGGGAATTCGTATT 122

QY 343 ValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleSerLeu 362
Db 123 ACCTCTCCCTATTTTACAACAGCTGTGAAGAGCTGCATGATGATCCTGCTCTT 182

QY 363 AlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly 382
Db 183 GCTGAAGGAATAGCAGTGGGAGAGCTTTGCAATGTTTCAAGAACTACAACATTGATGGG 242
```



```

Qy 383 AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402
Db 243 AACAAAGAAATGGTAGCATTTGGGACCATGACATTTGGTCTTCTTCTGTCTAT 302
Qy 403 LeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla 422
Db 303 CTCACACGGGGCCATCTCTCGATCAGCTGTAACTTCAATGCGGATGTAAACAGCA 362
Qy 423 AlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeu 442
Db 363 ATGTGGAACGTTGAATGGCCATTCAGTTATGTTTCATATTCATTTCTTAAACACCATTTG 422
Qy 443 PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu 462
Db 423 TTCCATTACACTCCCTCCCTGCTAGTGTCTCTGCAATTAATACTCTGCAATTCCTGGGCTC 482
Qy 463 IleAspTyrGluAlaIleIleHisLeuPheLysValAspLysPheAspPheValValCys 482
Db 483 ATGTATTCAGGGCTGCATTCATCTTTGGAAAGTCGATTAATTCGATTTTCGTTGTGTGT 542
Qy 483 MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle 502
Db 543 ATGGGTGCATATATCGGTGTGTCTTCGCGAGTGTGCGAGTTCGGATTCATAGCGGTT 602
Qy 503 ValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuGly 522
Db 603 GCTATCTCTGTCTCGAGATTCGCTCTCTGTTGTCGAGACCGAAGACCTTCATTTCTTGA 662
Qy 523 AsnIleProAsnSerValIleTyrArgAsnValGluHisTyr 536
Db 663 AACCTCCGAATTCACCATATATAGGAGTGTGAGCAATAT 704

```

```

RESULT 14
CK256297
LOCUS
DEFINITION
EST739934 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POC329 5' end, mRNA sequence.
CK256297
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

CK256297
783 bp mRNA linear EST 30-JUL-2004
EST739934 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POC329 5' end, mRNA sequence.
CK256297
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

REFERENCE
AUTHORS
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE
Generation of ESTs from potato callus tissue
JOURNAL
Unpublished (2003)
COMMENT
Other ESTs: EST739935
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ATT TAG GTG ACA CTA TAG.

```

```

FEATURES
Location/Qualifiers
1..783
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POC329"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/full_length"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 6.7e-108 Length: 783
Score: 1013.00 Matches: 193
Percent Similarity: 85.71% Conservative: 29
Best Local Similarity: 74.52% Mismatches: 35
Query Match: 29.15% Indels: 2
DB: 7 Gaps: 2

```

US-10-762-049-18 (1-680) x CK256297 (1-783)

```

Qy 24 MetGlySerValAspTyr---GluTyrProLeuGlyMetAsnAsnPheGluArg---Val 41
Db 5 ATGGGTAAACGAGAGATGATGATGATCCATCATCATGAAAGGGGAAACAGAAAAAA 64
Qy 42 HisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLys 61
Db 65 CACCAAGTTGAATTCACACCACCAACCATTTTTCAGTCACTTAAAAACACATGTAAAA 124
Qy 62 GluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLys 81
Db 125 GAAACATTTGTTTCAGATGATCCACTTAGGCAATTCAGCAATTCAGCAATTCAGCAAT 184
Qy 82 PheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPhe 101
Db 185 TTGAAACTTGGGGTTCAATATTTTCCCAATTTTCGAATGGGCGCTCGATATACGTTT 244
Qy 102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db 245 GATTTCTTAAATCTGATCTTATTTCTGGGATTACTATAGCCAGTCTTGTATTCCTCAG 304
Qy 122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPhe 141
Db 305 GGAATTTAGCTATGCTAAACTTGCCAAATTTGCCACCTATTCTTGGCCTCTATTCTAGTTT 364
Qy 142 IleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db 365 GTTCCAGCATTTAGTGACGAGTAATGGGTAGTTCAAGAGATTAGCAGTTTGAACAGTT 424
Qy 162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db 425 GCAGTTGTCATCTCTCTCAATGTTTAGGAGAGAAAGTAAATCCACATGATAAT 484
Qy 182 ProLysLeuTyrIleHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db 485 CCAACACTTTATCTTAATCTTGCAATTAACGCCACACATTTCTTCCCGGACTATTGAAGCC 544
Qy 202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db 545 GCATTTGGTATTTTTCAGGCTGGGATTCATTTGGATTTTCTATCTCATTCACAAATAGTG 604
Qy 222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db 605 GGATTCATGGGAGGACGACGACAGTTGTATCTCAACAAATTAAGGGAATCTTGGT 664
Qy 242 LeuGluHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThrGln 261
Db 665 CTTGACCATTTCACTCAGTCCACTGATATTAATTTCCGCTCTTGGTCCGCTTTTACCCAA 724
Qy 262 ThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280
Db 725 ACCGATGAGTGGCGATGGCAAGTGGGTGTTCTGTTTCTCTTCTTCTACCTG 781

```

RESULT 15

CF438329

LOCUS

DEFINITION

EST674674 normalized cDNA library of onion Allium cepa cDNA clone

ACAC64, mRNA sequence.

CF438329

ACCESSION

VERSION

CF438329.1 GI:34461019

KEYWORDS

EST

SOURCE

ORGANISM

CF438329 785 bp mRNA linear EST 04-SEP-2003

EST674674 normalized cDNA library of onion Allium cepa cDNA clone

ACAC64, mRNA sequence.

CF438329

ACCESSION

VERSION

CF438329.1 GI:34461019

KEYWORDS

EST

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae; Allium.

REFERENCE

1 (bases 1 to 785)

AUTHORS

Havey, M.J., Cheung, F., Van Aken, S., Uterback, T. and Town, C.D.
Expressed Sequence Tags from a normalized library of mixed onion

TITLES

tissues (Allium cepa)

JOURNAL

Unpublished (2003)

COMMENT

Contact: Havey MJ

Department of Horticulture

USDA-ARS and University of Wisconsin

1575 Linden Drive, Madison, WI 53706, USA

Tel: 608-262-1830

Fax: 608-262-4743

Email: mjhavey@facstaff.wisc.edu

TIGR sequence name ACAEC64TR. For more information:

http://haveylab.hort.wisc.edu

Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES

source

1..785

Location/Qualifiers

/organism="Allium cepa"

/mol_type="mRNA"

/cultivar="Red Creole (bulbs), unknown (callus), Ebano &

Texas Legend (roots)"

/db_xref="taxon:4679"

/clone="ACASC64"

/tissue type="Callus, roots, and young bulbs"

/clone_lib="normalized cDNA library of onion"

/note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site_1:

EcoRV (5'); Site_2: NotI (3'); Equal molar amounts of mRNA

from callus, roots, and young bulbs were combined to

synthesize the library. Normalization to enrich for

low-copy transcripts was performed by proprietary

techniques of Invitrogen."

ORIGIN

Alignment Scores:

Pred. No.:	2.28e-106	Length:	785
Score:	1000.00	Matches:	193
Percent Similarity:	82.29%	Conservative:	30
Best Local Similarity:	71.22%	Mismatches:	24
Query Match:	28.78%	Indels:	24
DB:	7	Gaps:	3

US-10-762-049-18 (1-680) x CF438329 (1-785)

QY	228	AlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHis	247
Db	2	GCAACTGTAGTTGTCTACAGCAGTGTGAAAAGTATTTGGGACTTCAGCATTTTACTACA	61
QY	248	GlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlu-----	264
Db	62	TCCACTGATCTCGTCTCTGTATGGAGTCTGTGTTGCAGTCTCACCAGGCATGCTGG	121
QY	265	-----TtpArgTtpGluSerAlaValLeuGlyCys	274
Db	122	CGTTGGAATATTCTAAGTAGTCCAGCAATGGAGATGGAGAGTGTGGTCTCTGGATGT	181
QY	275	ValPheIlePhePheLeuLeuSerThrArgTyrPheSerIlyLysArgProArgPhePhe	294
Db	182	GGATTTCTATCTTCTCTGTAGCACTCGCATTTTTCAGTAAAGAGGCAAAATTTATTT	241
QY	295	TtpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPhe	314
Db	242	TGGGTGCTGCGCAGCAGCTCCACTAACATCGGTCATCTTGGCAGTCTACTCGTCTACTTC	301
QY	315	ThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnPro	334
Db	302	ACCCAAGCAGAAAACCATGGTGTCAAGTATTGGTTATCTAAAGAAAGGGTTAAATCCA	361
QY	335	ProSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGly	354
Db	362	CCCTCTGTCAAAAAGCTGGTCTTCTCAGCACCTTATATTACGGTGTCTCTCAAAAGCTGGA	421

Search completed: August 31, 2005, 16:01:07
Job time : 5829 secs

QY	355	IleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMet	374
Db	422	ATCATCACTGGTATCATTTGCACCTTCTGAGGCATAGCAGTAGGAAGGAGCTTTTCAATG	481
QY	375	TyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnVal	394
Db	482	TTCAAGAATTACCATCGATCGAAACAAAGAGATGATCGCGCTTTGGGATGATGAACATG	541
QY	395	ValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsn	414
Db	542	GCCGGGTCCATCAGCATCATGCTACTTAAACACCGGACCAATCTCGAGGCACAGCTGTAAC	601
QY	415	TyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeu	434
Db	602	TACATGCAGGATGCAAGACGCAATGTCAATGTGGTATGCGCATAGCAGTATGATC	661
QY	435	ThrLeuLeuPheLeuThrProLeuPheHisTyrThrPro---LeuValValLeuSerAla	453
Db	662	ACGCTACTATTTTAAACACCGCTTTTCCACTACACTCCCTCTTGTGGTCCCTTGCCTCT	721
QY	454	IleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLys	473
Db	722	ATTATTATATCGCAATGCTTGTCTTATCGACTACGAGAGTGTCTCCACCTC	775
QY	474	ValAspLysPheAspPheValValCysMetSer	484
Db	776	-----TGCAATGTCG	784

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 11:07:04 ; Search time 425 Seconds
(without alignments)
9428.804 Million cell updates/sec

Title: US-10-762-049-17
Perfect score: 2449
Sequence: 1 gcacgagctagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2449	100.0	2449	4	US-09-720-317A-17
2	765.2	31.2	1981	4	US-09-720-317A-3
3	720.6	29.4	2067	4	US-09-720-317A-15
4	655.4	26.8	2279	4	US-09-720-317A-1
5	640.4	26.1	2311	4	US-09-720-317A-19
6	324	13.2	780	4	US-09-720-317A-7
7	216.2	8.8	1240	4	US-09-720-317A-5
8	205.6	8.4	2022	4	US-09-720-317A-21
9	118	4.8	493	4	US-09-720-317A-13
10	117.4	4.8	484	4	US-09-720-317A-9
11	90.4	3.7	2431	4	US-09-785-381-4
12	82.6	3.4	4927	4	US-09-949-016-5627
13	79.8	3.3	4113	4	US-09-785-381-2
14	77.6	3.2	2487	4	US-09-248-796A-6643
15	76.4	3.1	7218	1	US-08-232-463-14
16	64.4	2.6	2692	4	US-09-614-221A-222
17	63.8	2.6	2913	4	US-09-795-927-6
18	63.8	2.6	3749	4	US-09-795-927-8
19	58.2	2.4	510	4	US-09-720-317A-11
20	58	2.4	758	4	US-09-270-767-470
21	58	2.4	758	4	US-09-270-767-15752
22	57.8	2.4	3061	1	US-09-570-842-1
23	57.8	2.4	3061	1	US-08-700-576-1
24	56	2.3	1773	4	US-09-902-540-7746
25	56	2.3	4854	4	US-09-902-540-768
26	53.4	2.2	2832	4	US-09-949-016-4516
27	53.4	2.2	8774	4	US-09-949-016-16258

28	53.2	2.2	320	4	US-09-270-767-27504	Sequence 27504, A
29	53.2	2.2	904	4	US-09-270-767-11855	Sequence 11855, A
c 30	52.8	2.2	1116	4	US-09-252-991A-7460	Sequence 7460, Ap
31	52.8	2.2	1791	4	US-09-252-991A-7241	Sequence 7241, Ap
32	51	2.1	761	4	US-09-270-767-1826	Sequence 1826, Ap
33	51	2.1	761	4	US-09-270-767-17108	Sequence 17108, A
34	51	2.1	1767	4	US-09-252-991A-7184	Sequence 7184, Ap
35	47.4	1.9	1971	4	US-09-875-811-9	Sequence 9, Appl
36	47.4	1.9	1992	4	US-09-875-811-5	Sequence 1, Appl
37	47.4	1.9	2040	4	US-09-875-811-13	Sequence 13, Appl
38	47.4	1.9	2420	4	US-09-875-811-13	Sequence 14, Appl
39	47.4	1.9	7218	1	US-08-232-463-14	Sequence 22, Appl
c 40	45.4	1.9	1141	4	US-09-806-708B-22	Sequence 913, App
41	45.4	1.9	1143	4	US-09-540-236-913	Sequence 25, Appl
c 42	45.4	1.9	31147	4	US-09-596-002-25	Sequence 914, App
43	45.2	1.8	1466	4	US-09-620-312D-914	Sequence 5349, Ap
44	45.2	1.8	2858	4	US-09-949-016-5349	Sequence 1, Appl
45	45.2	1.8	2882	1	US-08-424-567-1	

ALIGNMENTS

RESULT 1
US-09-720-317A-17
; Sequence 17, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 2449
; TYPE: DNA
; ORGANISM: Glycine max
US-09-720-317A-17

Query Match	100.0%;	Score 2449;	DB 4;	Length 2449;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 2449;	Conservative 0;			
QY	1	GCACGAGCTAGCTCGCACATTAAAGTTATATACACATATTTGCTTGTAGAAATACTAT	60	
Db	1	GCACGAGCTAGCTCGCACATTAAAGTTATATACACATATTTGCTTGTAGAAATACTAT	60	
QY	61	TATTGAAGATATGGGAGGTAGATATGAGTACCCCTTTGGGCATGAACACTTTGAGAG	120	
Db	61	TATTGAAGATATGGGAGGTAGATATGAGTACCCCTTTGGGCATGAACACTTTGAGAG	120	
QY	121	AGTGACCAAGTCGAGGTTCCACCCGACGCGTTTTTCAAGTCTCTAAAGTACTCTTT	180	
Db	121	AGTGACCAAGTCGAGGTTCCACCCGACGCGTTTTTCAAGTCTCTAAAGTACTCTTT	180	
QY	181	GAAGGAGACTTTCTTCCCTGATGACCCCTTTGAGGCAGTTTCAAGAACAGCAGCTTCCAA	240	
Db	181	GAAGGAGACTTTCTTCCCTGATGACCCCTTTGAGGCAGTTTCAAGAACAGCAGCTTCCAA	240	
QY	241	GAAGTTCAAGTGTGGCCCTTCAGTCTTCTTCCCAATTTTGAATGGGCTCCCAATACAC	300	
Db	241	GAAGTTCAAGTGTGGCCCTTCAGTCTTCTTCCCAATTTTGAATGGGCTCCCAATACAC	300	
QY	301	CTTTTCAGTCTTGAAGCTGACCTCATAGCTGGCATCCATCGCTAGCTTGGCCATTC	360	
Db	301	CTTTTCAGTCTTGAAGCTGACCTCATAGCTGGCATCCATCGCTAGCTTGGCCATTC	360	

APPLICANT: Stephen M. Allen
APPLICANT: Saverio C. Falco
APPLICANT: Catherine J. Thorpe
TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
FILE REFERENCE: BB-1167
CURRENT APPLICATION NUMBER: US/09/720,317A
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/092,833
PRIOR FILING DATE: 14-07-1998
NUMBER OF SEQ ID NOS: 31
SOFTWARE: Microsoft Office 97
SEQ ID NO 3
LENGTH: 1981
TYPE: DNA
ORGANISM: Zea mays
US-09-720-317A-3

Query Match 31.2%; Score 765.2; DB 4; Length 1981;
Best Local Similarity 65.8%; Pred. No. 4.3e-202;
Matches 1112; Conservative 0; Mismatches 578; Indels 0; Gaps 0;

QY 313 GRAAGCTGACCTCATAGCTGGGCATCACCATCGCTAGCTTGGCCATTCCTCAGGGCATCAG 372
DB 7 GGAGTCGACCTGATCGCGGCATCACCATCGCCAGCCTCGCCATCCCGAGGGCATCAG 66

QY 373 TTATGCCAAGCTCGCCAAACCTCCCTCCAAATTCCTTGGACTATATTCGAGCTTTTATACCAC 432
DB 67 CTACGCCAAGCTCGCCAAACCTCGCGCGCTGCTCGACTCTACTCGAGCTTCGTGCGGC 126

QY 433 ATTGATTATGCAATGAGTGTAGCTCGAGGGATTTGGCAGTGGGGACGTGTGGCGTGG 492
DB 127 GCTGGTGTACGCTTGTAGTGGAGCTCCAAAGACCTGGCGGTGGGGACGTGTGGCGTGGC 186

QY 493 ATCGCTTCTGATGGGTTCGATGTTGAGTAATGCGCTTGATGCCAATGAAGACCCAAAGCT 552
DB 187 GTGCTGCTCATCAGCTCAATGCTCGGCAGCGAGGTGTGCGCCAGCGAATCCCGTGTCT 246

QY 553 TTACCTCCACCTCGCTTTTACAGCTACATTTATTTGCTGTGTGTTTTCAGGCTGCCCTTGGG 612
DB 247 CTACCTGCACCTCGCTTTTACCGCCACCTTCTTCGCGCGGTCTTTCAGGCTCGCTCGG 306

QY 613 TCTGTTAGGTTGGGTTGATGCTGTGATTTCTGTCAATGCAATCAATATAGGTTTCAT 672
DB 307 CCTCCTCAGGTTGGGCTTCATCGTGACCTGCTGTGCGACGCGATCGTGGGGTTCTAT 366

QY 673 GGGAGGACGACGACCGTGTGTCTGACGACCAATAAATCGATTCCTTGGCTTGGCTTGGACGA 732
DB 367 GGCGGCGCGGCGACGGTGTGTGCTGACGACGCTGAAGGGCATGCTGGGGCTGTGCTCCA 426

QY 733 TTTTACCCATGGAGCTGATATCATATCATGATGATGCTGCTGTTTTCACCCAAACTCATGA 792
DB 427 CTTTACCACTTCCACCGAGCTGCTTCCGTCATGGAATCCGTCTTCAGCCAGCACACCA 486

QY 793 GTGAGGTGGGAAAGTGTGTGTAGGATGTGTTCATTTTCTTCTCCTCTTAGCACAAG 852
DB 487 GTGGCGTGGGAGAGCTGCTGTGCGCTGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 546

QY 853 ATACTTCAGCAAAAAGCAGGCTTTTGTGGGTGTGACGATGCGCCCATTTGACGTC 912
DB 547 CTTTATCAGCAAGAGCGTCCCAAGCTGTTCTGGATCTCGCGGCGCGCGCTTGGACGTC 606

QY 913 CGTTATATTGGGAAGTCTCTTGGTGTATTTTCACTCACGCGGAGAACGCGTGTGTAAGT 972
DB 607 CGTGTGCTCGGAGGCTTCTGTGTACTCAGCGACGCTGAATACACACGCGATCGAAGT 666

QY 973 GATAGAGAACTGAAGAGGTTTGAATCCACATCACTCAAAATCTGATTTGTGTC 1032
DB 667 GATCGGTTACCTGAAGAAAGGCTGAATCCACGCTCGGTGACAAAGCTGCAATTTCTCACC 726

QY 1033 GCCTTACATGACTACAGCTGTCAAACTGGCATTTGCTGGCATCATATCACTTTCGGA 1092
DB 727 GGCCTACATGATGCTCGCGCTCAAGACTGGGATCATCCGCGGCTCATTTGCCCTCGCCGA 786

QY 1093 AGGAATAGCAGTAGGAAGAGCTTTGGCAATGTATAAAAATTAATAATTTGATGGCAACAA 1152
DB 787 AGGAATCGCGTGGGAGGAGCTTCGCCATGTTCAAGAACTACCAATGACGCGCAACAA 846

QY 1153 AGAGATGATAGCTATTGGGACCATGAACGTAGTGGTCTTTTCACTCTTGTACTCTAC 1212
DB 847 GGAGATGATCGCGATCGGACGATGAACGTCTCGGCTCGCTCAGCTGTGTCTACCTGAC 906

QY 1213 AACAGGACCATTTTCGCGTGTGAATTAACGCTGGATGCAAGACGACGAGCTTC 1272
DB 907 CACGGGCGCTTCTCGCGCTCGCGCTGAACTACAACGCGGCTGACAGACGCGCATGTC 966

QY 1273 CAACATTAATGTCATTTGCAATGTTGATTTTATTTCTTGACACCTTGTTCAC 1332
DB 967 GAACTGTGTCATGCTGGCGGTGATGCTCACGCTGCTGTTCTTCTGACGCGCTGTTCAC 1026

QY 1333 TTACACTCCCTGGTGGTCTATCATGCTATTTATCTGTAATCTGCAATGCTTGGACTCATAG 1392
DB 1027 CTACACGCGCTGGTGGTGTGCTGTCGCGGATCATCTGCTCCGCGATGCTGGGCTTGGTCCA 1086

QY 1393 TTATGAAGCAGCATCTCATTTTAAAGGTTGACAAATTTGACTTTGCTGTGATGAG 1452
DB 1087 CTTGCGGCGCGCTGCACTGTGCGGCTGCAAGGTGCACTTCTGCTGTGCGCTGCGCGG 1146

QY 1453 TGCATACATTTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTATGATTTGTAAT 1512
DB 1147 CGCGTACCTGGCGTGGTCTTTCGCGCAGCGTTCGAGGTGCGGCTGGTCTGCGCGTGC 1206

QY 1513 ATCTGTACTTGGGTACTTCTATTTTATGCAAGGCGCAAGGACATTCGTTTGGGCAACAT 1572
DB 1207 CTCCCTGTCTCGCGCTGCTGCTTTCGCGCGCGCGCGAGGACCAAGGTGCTCGGCAACAT 1266

QY 1573 TCCAAATCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATGTTCTCTCG 1632
DB 1267 CCGCGCGCACCTGCTGTACCGGAGATGGAACAGTACGCGCGCGCGCAGACGCTGCGCGG 1326

QY 1633 AATGTAATTTAGAGATTTGATGCAACATTTTATTTTGCCTTTCCTGCAATGCGAGCTATTTAGAGA 1692
DB 1327 CGTGTCTGTGCTGCGCTGCGCGCTGCTTCTTCTGCAACGCGCAGCTACCTGCGAGA 1386

QY 1693 AAGGATCACAGGTGGATGATGAAGAGAGAAAGAAATTAAGCTACAGGGGAGACTAG 1752
DB 1387 GAGGATCTCGCGGTGGATGACGACGAGGAGGAGCGCACCAAGAGCCAGGCGGAGATGGG 1446

QY 1753 TTTGCAATGTTTAAATTTGATGATGATGCTGCTTGGCAACATTTGATACAAGTGGAAATAG 1812
DB 1447 CGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506

QY 1813 TATGCTTGAAGAGGTGAAGAGATTTACAGAGAGAGAGAGCTACAGCTTGTGTTTGTCTCAA 1872
DB 1507 CATGCTGGACGAGCTCAACAGTCTTGGACAGGAGGAGGAAATGCAGATCGTGTGCGGAA 1566

QY 1873 TCCTGTAGTGAAGTGAAGAGAACTGAACAAATGCAAGTTTCCAAATCATTTAGGAA 1932
DB 1567 CCGCGGACGAGAGATCATGAAGAGCTGACAGCTTCAAGGTGCTTGGAGCAGATCGGCCA 1626

QY 1933 GAAATGAGTCTATCTGACTGTTTGAAGAGCGCGTTCGAGCATGCAACTTCAAATCTACGTCG 1992
DB 1627 CGAGTGGGTGTTCCCGACGCTGGGCGAGCGGCTGGGCTGCTGCGACTACGCTGCTGCACTC 1686

QY 1993 AAGCAAAACG 2002
DB 1687 GCACAAGCGC 1696

RESULT 3
US-09-720-317A-15
; Sequence 15, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167

; CURRENT APPLICATION NUMBER: US/09/720,317A

; CURRENT FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 14-07-1998

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 15

; LENGTH: 2067

; TYPE: DNA

; ORGANISM: Glycine max

; US-09-720-317A-15

Query Match 29.4%; Score 720.6; DB 4; Length 2067;
Best Local Similarity 62.2%; Pred. No. 1.1e-189;
Matches 1134; Conservative 0; Mismatches 689; Indels 0; Gaps 0;

QY	143	CGGCCACAGCGGTTTTCAGTCTCTAAAGTACTCTTTGAAGGAGACTTCTTCCCTGAT	202
DB	8	CCACACCGACACACTCCACAACTTAGGCACAGAGTCTCGGAAATCTTCTTCCAGAT	67
QY	203	GACCCCTTGGGAGTTCAAGACAGCCAGCTTCCAAAGTTCATGCTTGGCCCTCAG	262
DB	68	GACCCCTCTCACCGTTTCAAGAACCAACTCGCTTTAAAGATTTCTCTCGCACTTCAG	127
QY	263	TTCTCTTCCCAATTTTGAATGGGCTCCCAATACACCTTTTCAGTCTTTGAAGCTGAC	322
DB	128	TATCTCTTCCCAATTTTCAGCTGGGCCCACTCACTTACCTTCCGCTCTGAC	187
QY	323	CTCATAGCTGGGATACCAATCGCTAGCTTGGCCATTCCTCAGGGCAATGATGACCAAG	382
DB	188	CTCATCTGGGCTCACTTCCATTTGCGAGCTCGCCATTTCTCAGGGAATCAGTTATGCCAAG	247
QY	383	CTCGCAACTCCCTCCATTTCTTGACTATATTCAGCTTTATACCACTTGAATTTAT	442
DB	248	CTTGCCAACTGCGCACTATTTCTTGATATATTCAGATTTTGTTCCTCCCATTTGATATAC	307
QY	443	GCGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGAGCTGTGGCGGTTGGATCGCTTCG	502
DB	308	TCGCTGCTTGGAGTTCTAGACATCTTGGTGTGGACCTGTTCCATTCGGCTTTGGTC	367
QY	503	ATGGGTTCCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	562
DB	368	ATGGGATCAATGTTAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	427
QY	563	CTGGCTTTCACAGCTACATTTTGTGCTGTTGTTTTCAGCTGCTGCTGCTGCTGCTGCTGCT	622
DB	428	TTGGCTTTCACCGCCACTTCTTGTGCTGTTGATGATGATGATGATGATGATGATGATGATGAT	487
QY	623	TTGGGCTTGCATCGTGATTTTCTGTCACTGCAACCATTAATAGGTTTCATGGGAGGACA	682
DB	488	CTAGGCTCTGTAATTCATTTCTGTGCAAGGCAACGCTGTTGGATTTCACGCGGTGCT	547
QY	683	GCCACGGTGGTGTCTGCGCAACTTAAATCGATTTCTTGGCCCTTGGCAATTTACCCAT	742
DB	548	GCCATTAATTTGTCTACTGCGAGCTGAAAGGTTTACTTTGGAATAGTGCATTTACCCAG	607
QY	743	GGAGCTGATATCATATCATGATGCTGCTGTTTTCACCCCAACTCATGAGTGGAGTGG	802
DB	608	AGATGCAAAATATTCAGTAAGCATCTCTGTTTTCAGCAAGACACCGAGTGGTCATGG	667
QY	803	GAAAGTGTGTGTAGGATGTGTTCTTCAATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	862
DB	668	CAAAACCATTTCTTGGGATTCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	727
QY	863	AAAAACGACCAAGGTTTTTTTGGGTGTGCAAGATGCGGCCATTTGACGTCGGTTATATG	922
DB	728	TTGAGGAAACCAAAACTATTTCTGGGTTTCAGCAGCTGCCCCCATTTGACATCAGTTATCTG	787
QY	923	GGAAGTCTCTGTTGTTTATTTTCACTCAGCGCGAAGCAGCGGTGTTCAAGTGTATGAGAA	982
DB	788	TCAACCAATTTAGTCTTTCTTCTGAGAAATAGACTCATCAATTTTCAGTTATTTGGGCAC	847

RESULT 4

US-09-720-317A-1

; Sequence 1, Application US/09720317A

Patent No. 6696292
GENERAL INFORMATION:
APPLICANT: Stephen M. Allen
APPLICANT: Saverio C. Falco
APPLICANT: Catherine J. Thorpe
TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
FILE REFERENCE: BB-1167
CURRENT APPLICATION NUMBER: US/09/720,317A
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/092,833
PRIOR FILING DATE: 14-07-1998
NUMBER OF SEQ ID NOS: 31
SOFTWARE: Microsoft Office 97
SEQ ID NO 1
LENGTH: 2279
TYPE: DNA
ORGANISM: Zea mays
US-09-720-317A-1

Query Match 26.8%; Score 655.4; DB 4; Length 2279;
Best Local Similarity 60.4%; Pred. No. 1.6e-171;
Matches 1120; Conservative 0; Mismatches 726; Indels 9; Gaps 2;

QY 121 AGTGCAACAGTCGAGGTTCCACGGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTT 180
DB 188 AGTGCAACAGGTGGCGCGCCACCGCGCGGAGCACGGCGAGCAAGATGAAGGTGAGGTT 247
QY 181 GAAGGAGACTTCTTCCCTGATGACACCTTTGAGGCGAGTTCAAGAACAGCCAGCTTCCAA 240
DB 248 GAAGGAGACTTCTTCCCGACAGCCGTTTCCGGGCGTTTCAAGGGGAGCGCGCGGGAC 307
QY 241 GAAGTTTCATCTTGGCTTCAGTTCTTCTTCCCATTTTTCGAATGGGCTCCCAAAATACAC 300
DB 308 GCAGTGCTCATGGCGGTGAGTACCTTCTTCCCATCTTCCCATCTGGACTGGGTGCGAGTACTC 367
QY 301 CTTTTCAGTTCTTGAAGCTGACCTCATPAGCTGGCATCACCATCGTAGCTTGGGCCATTC 360
DB 368 CTTTTCAGTTCTTCAAGTCCGACCTCGTGGCGGCTCACCATTGGCCAGCTCGCCATTTC 427
QY 361 TCAGGGATCAGTTATGCGAAGCTCGGCAACCTCCCTCCAAATCTTTGGAGTATATTCGAG 420
DB 428 TCAGGGATCAGTTATGCGAAGCTGCGAAGCTTGGCTCCCATTAATCGGGCTGTATTCGAG 487
QY 421 CTTTATACACCAATTCATTTATCGGATGAGGTGAGTCTGAGGATTTGGCAGTGGGAC 480
DB 488 CTTTTCGCGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 547
QY 481 TGTGGCGGTTGGATCGCTTCTGATGGGTTGATGATGATGATGATGATGATGATGATGATG 540
DB 548 GGTGTGATCTGCTGCTGATCATGATGATGATGATGATGATGATGATGATGATGATGATG 607
QY 541 AGACCCAAAGCTTTACTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 608 GGAGCGGACGCTTCTGCGAGTGGGCTTCACTCCACCGCTTTCGCGGGGCTGGTGCA 667
QY 601 GCGTCTTGGGCTGTTAGGTTGGGTTGATGATGATGATGATGATGATGATGATGATGATG 660
DB 668 GCGCTTCTGCGGATCTTCAAGCTCGGCTTCTGATGATGATGATGATGATGATGATGATG 727
QY 661 AATAGGCTTCATGGGAGGAGCAGCCAGGTTGGTGTGCTGCGAGCAATCAAAATTCGATTC 720
DB 728 GGTGGGTTTCATGGCGCGCGCCCATCATCGTGGCGCTGCGAGCACTCAAGGGGCTGCT 787
QY 721 TGGCTTTGAGCTTCAACCATGAGAGTATATATATATATATATATATATATATATATATAT 780
DB 788 GGGCATCGTCCACTTCAACACAGAGTGGGATCGTCCAGTCAATGGGCTCGCTCTTCCA 847
QY 781 CCAACTCATGAGTGGGAGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 848 CCACACAGGAGTGGTGGGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 907
QY 841 CCTTAGCACAGATACTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTTGTCAGCAATGGC 900

DB 908 GCTGTGCGGAGGCATGTGAGCATCAGATGGCCAAAGCTTTTCTGGGTTTTCGCGCTGCGC 967
QY 901 GCCATTGACGTCCGTTATATTGGGAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGCA 960
DB 968 GCCCTTGGCATCGGTCAACATCTGACGCTGCTTGTTCCTTCCAAAGCTCAGAACCA 1027
QY 961 CGGTGTTGAAGTGATAGGAGAACTCAAGAAAGGTTTGAATCCACCATCACTCAAAATCT 1020
DB 1028 TGGCATCAGCATCATTTGGGCAGCTCAAGTGGCGCTGAATCGCCCTCGTGGGACAAGCT 1087
QY 1021 GGTATTTGTGTCGCTTACATGACTACAGCTGTCAAACTGGCATTTGTGTGGCATCAT 1080
DB 1088 CCTGTTTGACACGGGCTATTTAGGCTCACCATGAAGACTGGCTTTGTCCCGAATCAT 1147
QY 1081 ATCACTTGGGAAGGAATAGCAGTAGGAGAAAGCTTTGCAATGTATAAAAAATTTACAATAT 1140
DB 1148 CTCACCTGAGGAAGGAATAGCGGTTGGTAGAACATTTGCTCCTCCTCACTCAAGGACTACAGAT 1207
QY 1141 TGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGAGTAGTTGTTTCTTCACTC 1200
DB 1208 AGATGGMAACAAGGAGATGATGCGCCATAGGGTTGATGAATGTTGTTGGGCTCTGCACATC 1267
QY 1201 TTGCTACCTCAACACAGGACCAATTTTGGGTTTCGGCTGTGAACATATAACGCTGGATGCAA 1260
DB 1268 ATGCTACGTAAACACAGGTGCGTTCTCCGCTCTGCTGTAAACCAACACCGCGCTGCAA 1327
QY 1261 GACAGCAGCTTCCAAACATTAATATGCTTCACTTGCAGTAAATGTTGACATGTTTCTCTGAC 1320
DB 1328 GACTGCGATGTCAAACGATGATGCGGCTGACTGTGATGGTCAAGCTGCTGTTCTCAT 1387
QY 1321 ACCCTTGTTCATTAACATCCCTCGTGGTGTCTATCAGCTATTATGCTATCTGCAATGCT 1380
DB 1388 GCCACTGTTCTGTACACACCAACGTTGTCTCGGAGCGATCATCATCGCGGCTGAT 1447
QY 1381 TGGACTCATPAGATTAAGAGAGCCATCCATCTATTTAAGTTGAACAAATTTGACTTTGT 1440
DB 1448 CGGCTGATCGATTTCCCGCGGTGTACCACTCTGGAAGATGGAACAGATGATTTCT 1507
QY 1441 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1508 GGTGTGCTTTCGCGGTTTTCGCGGCTCATCTTCATCTCAGTCCAAAGAGGCTTCGCGAT 1567
QY 1501 AGCTATTGTAATATCTGTACTTTCGGGTACTTCTATTTATTTGCAAGGCCAAGGACATTCGT 1560
DB 1568 AGCGGTTGATATCTATATTTAGGGTTGATGAGATCAACAGGCCGAGATGATGAT 1627
QY 1561 TTTGGGCAACATTCCAAATTCGTGATATACCGAAATTTGTGAGACTATCAAAATGCAAA 1620
DB 1628 TCAAGGGAAACATCAAGGGGACTGATATTTACAGAGACCTGCTCATCACTACAGGAGGCCCA 1687
QY 1621 ACATGTTCTGGAATGCTAAATTTAGAGATTCATGCAACCAATTTACTTTTGGCAATGCCAG 1680
DB 1688 AAGAGTTTCTGGGTTTCTGATCTTGGCCATTTGAAGCACCGATAAACTTCGCCAATCCAA 1747
QY 1681 CTATTTAGAGAAAGGATCAAGGTGGATTCATGAAGAAAGAAAGAAATTTAAAGCTAC 1740
DB 1748 CTACCTGATGAAGAGATTAAGAGATGAT-----AGAGGAAGATCTTTTGAACAGCA 1801
QY 1741 AGGGAGACTAGTTTTCAGTATGTTTATATTTGATATGAGTGTGTGGAACCAATTGATAC 1800
DB 1802 TAAACATCTGAACCTCCATTTCTAATCTTGGATCTGTGAGTCTTCTGCAATTTGACAC 1861
QY 1801 AAGTGAATAGTATGCTTGAAGAGGTGAAGAGATTACAGAGAGAGAGAGACTACAGCT 1860
DB 1862 AAGTGGCATAGCGTTTCTCATTTGACATAAAGAAATCAATAGAGAAACGTTGGTCTGAGCT 1921
QY 1861 TGTGTTTGGTCAATCTGTAAAGTGAAGTGAAGAAATTTGAAGAAATTTGAAGAAATTTGAAG 1917
DB 1922 TGTGTTTGTCAATCCAACTGGGAGAGTCAATGGAGAAATATCAACGTTGCAACAGGCTGA 1981
QY 1918 AAATCAATTTAGGGAAGAAATGGAATCTATCTGTCTGTGAAGAGGCGGTTGGAGCA 1972
DB 1982 AAATCAATTTAGGCGCAGATTGCTGTATCTGACCACCTGGCGAAGCAATCGCTTCA 2036

RESULT 5

US-09-720-317A-19

; Sequence 19, Application US/09720317A

; Patent No. 6696292

; GENERAL INFORMATION:

; APPLICANT: Stephen M. Allen

; APPLICANT: Saverio C. Falco

; APPLICANT: Catherine J. Thorpe

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167

; CURRENT APPLICATION NUMBER: US/09/720,317A

; CURRENT FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 14-07-1998

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 19

; LENGTH: 2311

; TYPE: DNA

; ORGANISM: Triticum aestivum

US-09-720-317A-19

Query Match 26.1%; Score 640.4; DB 4; Length 2311;

Best Local Similarity 59.4%; Pred. No. 2.3e-167;

Matches 1104; Conservative 0; Mismatches 751; Indels 3; Gaps 1;

QY	122	GTGCACCAAGTCGAGGTTCCACGCCACACAGCCGGTTTTTCAAGTCTCTAAAGTACTTTTG	181
DB	187	GTGTACAAAGTGGCTATCCCTCCGAAGAATTTGGCCACAGAGTTTACAGAAACATTG	246
QY	182	AGGAGACTTTCTCCCTGATGACCCCTTTAGGAGGTTTCAAGACAGGACGCTTCCAG	241
DB	247	AGGGAGACTTTCTTCCACGACAAACCGCTGCGTCAATTAAGGGCAATCCGGACCGAGG	306
QY	242	AAATTCACTGGTGGCTTCAGTTCTTCTCCCATTTTCGAATGGCTCCCAAAATACACC	301
DB	307	AGTTTCATGAGGGCTGGAGTTCTTTGTTCCATATTTTGGTGGGTAGGATTTACAGT	366
QY	302	TTTTCAGTTCTTGAAGCTGACCTCATPAGCTGGCATCACCATCGCTAGCTTGGCCATTCCT	361
DB	367	CTCAACAAGTTCAAAGCGATCTGATTCGGGATTGACCATCGCAAGTCTCTGTATTCCT	426
QY	362	CAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTTGGACTATATTGAGC	421
DB	427	CAGGACATTTGGCTATTTCGAAGCTTGTCTAATCTGGATCCGCAATATGGGCTTTACTCCAG	486
QY	422	TTTATACCAACCATTTGATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACT	481
DB	487	TTCAATTCCTCCATTTGATCTATGCTGCAATGGGTAGCTCAAGGATATAGCGATTTGTCCA	546
QY	482	GTGGCGGTTTGGATTCGCTTCTGATGGGTTTCGATGTTTGGTAAATGCCGTTGATCCCAATGAA	541
DB	547	GTTCGTGTGGTTCTCTTTTGTATAGTTTCACTTCTACAGCTGAGGTGACCATGTCAA	606
QY	542	GACCCAAAGCTTTCCTCCACTGGCTTTTCAAGCTATACATATTATTCCTGGTGTTCAG	601
DB	607	AAACAAGGAGAAATACATCGCCCTCGCTTTTCAAGCAACCTTCTTCGCTGGTATCACTCAA	666
QY	602	GCTGCCTTGGGTCTGTTTGGTGGGTTGATCGTGGATTTCTGTCAATGCAATGCAATATA	661
DB	667	GCAGCCTTAGGATTTCTAAGGTTAGGATTCCTTATAGAGTTCTTTGTCGATGCTGGGAT	726
QY	662	ATAGGGTTTATGGGAGGACGACCGGTGGTGTGCTGCAGCAATATAAATCGATTCCT	721
DB	727	GTTCGATTTATGGGGGAGCTGCCATTAATTTGCTGCAGCAGCTGAAATACGTTG	786
QY	722	GGCTTTGAGCAATTTCAACCATGGAGCTGATATCATATCAGTATGCTGCTGTTCAC	781
DB	787	GGCATCGCAAACTTTTACAAGGAAACCGACATAGTTTCTGTATGGAATCTGTCTGGAGA	846
QY	782	CAAACTCAT---GAGTGGAGGTGGGAAAGTGCTGTGTAGGATGTCTTCATTTTCTTC	838

DB	847	TCAGTTTCATACCGGGTGGAACTGGCAGACAATTTGATTTGGCGTATCTTTCTCGTTTC	906
QY	839	CTCCTTAGCAAGATATCTTTCAGCAAAAGACCAAGGTTTTTTTGGTGTACAGCAATG	898
DB	907	CTTCTGTTTGGAAAGTACATCGGAAAGAAAGAAAGGAGCTTTTTCTGGGTCCAGCTATT	966
QY	899	CGGCATTTGACGTCGCTTATATTGGGAAGTCTCTTTGGTTTTTATTTTCACTCAACGCGAGAAG	958
DB	967	GCTCTATAATTTTCACTGATTTCTAGCAACAATTTTGTATATCAATTTCTCGTGGCGACAAG	1026
QY	959	CACGGTGTGAAGTGAATAGGAACTGAAGAAAGGTTTGAATCCACCATCTACAAAT	1018
DB	1027	CAAGGAGTTTCAATAGTAGAAGCAATTCACACAGGGAATCAACCCATCATCAGTACACAAG	1086
QY	1019	CTGGTATTGTCGCTTACATGACTACAGCTCTCAAACTGCGCAATTTGCTGTTGGCATC	1078
DB	1087	ATTTATTTTCAACCGGCCCATTTTGTGCAAAAGGTTTCAAGATCGGTGTTGTTGGGCGATA	1146
QY	1079	ATATCACTTGGGAAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATATAAAATTTACAAT	1138
DB	1147	GTGCTTTGACAGAAAGCTGTAGCTATTGGAAGGACATTTGCTGTATGAAGGACTACCAG	1206
QY	1139	ATTGATGGCAACAAAGAGATGATAGCTATTGGGACCAATGAACGTAAGTGTGTTTCACT	1198
DB	1207	TTAGATGAAACAAAGGAGATGGTAGCACTTGGAAACCAATGAACATAGTAGGCTCAATGACA	1266
QY	1199	TCCTGTACTCTACACACAGGACCAATTTTCGGTGTGGCTGTGAACTATAACGCTGGATGC	1258
DB	1267	TCCTGTATGTCAACAACAGGTTCTTCTCAGCTTGGCAGTTAACTTCTATGGCTGGCTGC	1326
QY	1259	AAGCAGCAGCTTCCAACTTAATGATGCACTTGCAGTAATGTTGACATTTGTTATCTCTG	1318
DB	1327	AAAGCTCTGTATCCATGTGCTGTTATGTCAGTAGTGGTCTTCTTACCTTGTGCTCATC	1386
QY	1319	ACACCTTGTTCATTTACACTCCCTGGTGGTGTATCAGCTATTATCGTATCTGCAATG	1378
DB	1387	ACACCGCTATTCAAATAFACACCGAATGCAATCCTAGGGTCGATCATTTATTTCTGCGGTG	1446
QY	1379	CTTGGACTAGATTAAGAAGCAGCCATCATCTATTTAAGGTTGACAAATTTGACATTT	1438
DB	1447	ATCGCCCTTGTGGACTACGAAAGCAATCTCATCTGGAAGTTGACAAATTTGACATTT	1506
QY	1439	GTGGTGCATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTC	1498
DB	1507	ATTGCTTGCATGGAGCTTTTTTGGTGTGTTTTTGTATCCGTTGAGATTTGGCTCTTG	1566
QY	1499	ATAGCTATTGTAATCTGTACTTTGGGTACTTCTTATTTATGCAAGCCCAAGGACATTC	1558
DB	1567	ATTGCTGTAGCAATCTCATTTGCCAAATACTTCTTCAAGTAAACAAGGCCAAGGACAGCC	1626
QY	1559	GTTTTGGGCAACATTCGAAATTTGTGTATATACCGAATTTGAGCACTATCAAAATGCA	1618
DB	1627	CTACTTGGAAACCTTCCCGGACCACTATATACCGGAACTATCAGCCAGTATCCGAAGA	1686
QY	1619	AAACATGTTCTCGAATCTAATTTCTAGAGATTGATGCAACCAATTTACTTTGCCAATGCC	1678
DB	1687	AAACTTACTCTCTGGGTGGTGTGAGGGTTGATCTGCTATTTATTTTTCCTCAACTCT	1746
QY	1679	AGCTATTTAAGAGAAAGGATCAAGGTGGATTTGATGAAGAGAGAAAGAAATTTAAAGCT	1738
DB	1747	AAATACGTCGGAAGAAAGATTTCTTAGGTGGCTGACAGACGAAGAAAGACAGAGCTAAAGCA	1806
QY	1739	ACAGGGGAGACTAGTTTTCAGTATGTTTATTAATTAATGATGATGCTGTTGGAAACATGAT	1798
DB	1807	GTGGGATTTGCCCTTAAATCAGTTTCTCTGATTTGGAAATGTGCGCGGTCTATCGACATCGAT	1866
QY	1799	ACAACTGGAATAGTATGCTTCAAGGCTGAAGAGATTTACAGAGAGAGAGAGCTACAG	1858
DB	1867	ACAAAGCGGATACATGCTCTTGAAGATCTATACAAAGATCTTTCAGAAAAGATATGAG	1926
QY	1859	CTTGTGTTTGGTCAATCTCTGTAAGTGAAGTGAAGAAACTGAAACAAATCGAAGTTC	1918

Db 1927 CTCATTCTGCGAATCTCGTTCGTCGTCATAGAAAAAAGTCGAAGCGTGAAGCTCAAC 1986
QY 1919 AATCATTTAGGGAAGAAATGATCTATCTGACTGTTGAAAGAGCGCGTTGAGCATGCA 1976
Db 1987 GAGCATTGGAGGACGACATATATCTCTCGCGCTCTCTGACGCTGTGCGATTCTGTA 2044

RESULT 6
US-09-720-317A-7
; Sequence 7, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Helianthus tuberosus
US-09-720-317A-7

Query Match 13.2%; Score 324; DB 4; Length 780;
Best Local Similarity 70.6%; Pred. No. 1.2e-79;
Matches 432; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
QY 1400 GCAGCCATCCATTAATTAAGGTTGACAAATTTGACTTTGTGTGTCATGATGTCATAC 1459
Db 17 GCGCGGATTCACCTCTGGACACTAGACAAATTCGACTTTGTGTATGATGATGTCATAC 76
QY 1460 ATTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTATTGTAAATCTGTA 1519
Db 77 TTGGTGTGTTCTTTGGAGTGTGAAATTTGGATTTAGTTATCGCGTGCATTTGCTTG 136
QY 1520 CTTGGGTACTCTATTTATTTGCAAGCCCAAGACATTCGTTTGGGCAACATTCOAAT 1579
Db 137 CTTAGGCTACTCTATTTGTCTCGAGGCCAAGAACATCGACGCTAGTCTCATACCGAT 196
QY 1580 TCTGTGATATACCGAATGTTGAGCACTATCAAAATGCAAAACATGTTCTGGAATGCTA 1639
Db 197 TCCACTATTTATAGAAGTATGGATCAATACCAAAATGCGAAAAGCGTTCCAGGAATCTTG 256
QY 1640 ATTCTAGAGATTGATGCACCAATTTACTTTTGGCAATGCCAGCTATTTAAGAGAAAGATC 1699
Db 257 ATACTTCAATCGAAGCACCATTATTTCTTTGCTACTCTAGCTATTGAGGGAAGATTT 316
QY 1700 ACAAGTGGATTGATGAAGAGAAGAAATTTAAAGCTACAGGGGAGACTAGTTTTCAG 1759
Db 317 GTCAGATGGGTTGATGAAGGGAAGATAGTTTGAAGTCTTTAAAGGAGAATGACTTTGCAA 376
QY 1760 TATGTTATATTTGATGATGCTGTTGGAACATTTGATACAGTGGAAATAGTATGCTT 1819
Db 377 TATGTCATTTCTGATGATGCTGTTGGAATATTTGATACAGTGGGATAAATGCTT 436
QY 1820 GAAGAGGTGAAGAAGATTACAGAGAGAAGAGCTACAGCTCTGTTTGTTCATCTGTA 1879
Db 437 GGAGAAGTTAAAGGTTATGGAAGAAGAGGGCTTAAGTTGTTTTAGGAAATCCGGGC 496
QY 1880 AGTGAAGTGAAGAAACTGAAACAAATCGAAGTTCCAAATCAITTTAGGGAAGAAATGG 1939
Db 497 GGAGAGTAATAAGAAGATGAACAAGCGAAGTTGATAGAGGTGATCGGCGAAGATGG 556
QY 1940 ATCTATCTGACTTTGAAAGGCGGTTGGAGCATGCAACTTCAATCTGCTGCAAGCAAA 1999
Db 557 ATATATCTAACAGTGGGAGAAGCGGTTGGAGCGTGCAACTTTATGCTTCTATCTTACAAG 616

QY 2000 ACGAACCCCAAG 2011
Db 617 AACGCCGAAAG 628
RESULT 7
US-09-720-317A-5
; Sequence 5, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-5

Query Match 8.8%; Score 216.2; DB 4; Length 1240;
Best Local Similarity 53.4%; Pred. No. 1.4e-49;
Matches 477; Conservative 0; Mismatches 413; Indels 3; Gaps 1;
QY 1085 CTTGGGGAAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATAAAAAATTACAATATTGAT 1144
Db 9 CTCACGGAAGCTATCGCGTTGCGCGATCTTTGCGCTCCGTAAGAGGGTACAGACTCGAC 68
QY 1145 GGCACAAAGAGATGATAGCTATTGGGACCATGAAGAGTAGTTGGTTCTTTTCACTTTCG 1204
Db 69 GGCACAAAGAGATGCTGGCCATGGGTTCTCCAAAGTTGCTGGTTCTCTGCTCTCGTGC 128
QY 1205 TACCTCACAAGAGGACCAATTTTGGGTTGGCTGTGAACATAAAGCTGATGCAAGACA 1264
Db 129 TATGTGGCAACAGGTTGCTTCTCCGAAAGCGGAGTGAACCTTCAGCGGGGGCGAGTCG 188
QY 1265 GCAGCTTCCAACTATTAATGTCACCTGCAAGTAATGTTGACATTTGTTATTCCTGACACC 1324
Db 189 ACGTTTCAACATGTCATGTCATCACCGTTCGTCACCTGGAGCTGTTTCATGAAG 248
QY 1325 TTGTTCCATTACACTCCCTGGTGGTGTATCAGCTATTATCGTATCTGCAATGCTTGA 1384
Db 249 CTCCTCTACTACACGCCATGGCGTCTCGCTCCATCATCTCTGCGGCTCTTCCGGGA 308
QY 1385 CTCATAGATTATGAAGCAGCCATCCATCTATTTAAGTTGACAAATTTGACTTTGGTG 1444
Db 309 CTGATCGCATCAAGAGGCGCTGCAAGATATGGAAGATCGACAAGATGATTTCTTCACC 368
QY 1445 TGCATGAGTGACATACATTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCT 1504
Db 369 TGCTTGGTGGCTTTGTTGGCGTCTGTTGGGTGGTGGAGATTGGGGCTTCGAGTTGCA 428
QY 1505 ATTGTAATATCTGTAATCTGGGTAATCTTATTTATGGAAGGCCAAGGACATTCGTTTGG 1564
Db 429 CTTGGCATTTCTCTCGAAAGATCATACAGTCGCTTCGGGCTTCAGGTGGAGATCTT 488
QY 1565 GGCACATTCCAAATCTGTGATATACCGAATGTTGACACTATCAAAATGCAAAACAT 1624
Db 489 GCGAGGCTACAAAGGACAGATATCTTCTGACGCTCAGCAGTACCTCTGAGCTTGCCTA 548
QY 1625 GTTCTCGGAATGCTAATCTAGAGATTGATGACCAATTT---TACTTTGCCAATGCCAGC 1681
Db 549 ACTCGAGCTGTACTGCTTATCGGCTCGACATCTCTTCTCTGCTTCTATCAACGCCACT 608
QY 1682 TATTAAAGAAAGGATCACAGGTGGATTGATGAAGAGAAGAAAGAAATTAAGACTACA 1741

Db 609 TCCGTCAAAGGATCACAGATGGGTTTGGGAAGAGTGGAGACCTCAAATGAAAA 668
Qy 1742 GGGGAGACTAGTTTGCAGTATGTTAATTGATATAGTGTCTTGGAAACATTGATACA 1801
Db 669 GCGAGGGAGAGGATACAAAGCAGTTGTCTTGTATATGTCAAGTGTGGTAAACATCGACACT 728
Qy 1802 AGTGGAAATAGTATGCTTCAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTT 1861
Db 729 TCAGGACTCAGCTGCACTGGAAGAAATACAAAGGAGTTGGTGTCTCTTGGCTTACAGATG 788
Qy 1862 GTTTTGGTCAATCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAAAT 1921
Db 789 GCTATAGCCAGTCGSGATGGAAGCAGTTCAGAAGATGAAAGTGTCAAGGTGTGGAC 848
Qy 1922 CATTTAGGAAGAATGGATCTATCTGACTGTGTAAGAGCCGTTGGAGCATG 1974
Db 849 AGGGTAGGACAGGACTGGATCTTCATGACAGTAGGTGAAGCGTGGAGGCCGTG 901

RESULT 8

US-09-720-317A-21
; Sequence 21, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: *Triticum aestivum*
US-09-720-317A-21

Query Match 8.4%; Score 205.6; DB 4; Length 2022;
Best Local Similarity 47.0%; Pred. No. 1.7e-46;
Matches 791; Conservative 0; Mismatches 869; Indels 24; Gaps 4;
Qy 306 AGTTCTTGAAGCTGACCTCATAGTGGCATCACCATCGTAGTGGCCATTCCTCAGG 365
Db 103 AGGACTTCCAGGCCGACCTCGCGCGGCATCACTGTGCGCGTCATGCTTGTGCGCTCAGG 162
Qy 366 GATCAAGTTATGCCAAGCTCGCCAACTCGCTCCAAATTTCTTGGACTATATTTCGAGCTTTA 425
Db 163 CAATGTCAATATGCAAAAGCTGGCTGGGCTTTCAACCAATTTATGGGCTCTACACAGGCTTTG 222
Qy 426 TACCACATTTGATTTATGCGATGATGGTAGCTCGAGGAGATTGGCAGTGGGACTGTGG 485
Db 223 TCCACTATTGCTACGCGAATTTTGGTCTCAGCAATATGACAGATGAGTCCAGTGG 282
Qy 486 CGGTGGATCGCTTCTGATGGTTCGATGTGAGTAATGCCGTGTGATCCCAATGAAGACC 545
Db 283 CACTTGTCTCTGCTAGTGTCCAAATGTTCTTGGGGGTATAGTTAATTCATCTAGTG--- 339
Qy 546 CAAAGCTTTACCTCCAGCTGGCTTTTACAGCTACATATTGCTGGTGTGTTTTTCAGGCTG 605
Db 340 ---AGCTGTACACGGAATTAGCCATATTATGCGAATTCATGTTGGAACTACTGGAATGCT 396
Qy 606 CTTTGGTCTGTTTACGTTGGGTTGATCGTGATTTTCTGTACATGCAACCATAAATAG 665
Db 397 TGAATGCAATGCTAAGACTTGGCTTGGCTTATTCGTTTCATAGCCATTCTGTAATATCTG 456
Qy 666 GGTTCATGGAGGAGCAGCCAGGTGTGTCTGCAGCAACTTAAATCGATTCCTTGGCC 725
Db 457 GATTCACTACAGCTTCGGCCATCGTAATTGGTTGTTGCCCAATCAAGTATTTCTTGGG--- 514

Qy 726 TTGAGCAATTTCAACCATGGAGCTGATATCATATCAGTGAATCGGCTCTGTGTTTCAACCCAAA 785
Db 515 -TTACAGTGTATCAAGAAGTAGCAAAATTAACCACTTATTGAGAGTATAAATTGCTGAA 573
Qy 786 CTATAGGTGAGGTGGGAAAGTGTGTGTTAGAGATGTCTTCTCATTTTCTTCTCCTCTTA 845
Db 574 TAGATCAGTTCTCTCGCTCCATTTGTTAATGGGATCAGCGTTTCTTGTATTCTTCTTAA 633
Qy 846 GCACAGATATCTCAGCAAAAAACGACCAAGTTTTTTTTGGGTGTCAGCAATGGGCCAT 905
Db 634 TAATGAAAAAGCTAGGAAAAACAAATAAAATTTACGTTTCTTGAGAGCTTCTGTGTCAC 693
Qy 906 TGAGCTCGTTTATATTTGGAAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAACGCGTG 965
Db 694 TAAAGCTGTGTTCTTGGAAATTTGTTGTGAAAAATTTTCGGTCC-----AAGTGCA 747
Qy 966 TTGAAGTGAATAGGAGAACTGAAGAAGGTTTGAATCCACATCACTCAAAATCTGGTAT 1025
Db 748 TATCAGTGGTAGGTGAATAACCGCAAGGCCCTTCCAGGTTTCTCCATTTCTCGAGGATTTG 807
Qy 1026 TTGTGTCGCTTACATGACTACAGCTGTCAAACTGGCAATGCTGTTGGCATCATATCAC 1085
Db 808 AACATCTGATGTCCCTAATGCGCAACTGCAATCTTATCAGTGGTTG-----CTA 858
Qy 1086 TTGCGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTTACAATATTTGATG 1145
Db 859 TTTTGGAGTCTGTGGGATTGCTAAAGCGTTAGCTGCGAAGAAATGGTTATGAGTTGGACT 918
Qy 1146 GCAACAAAGAGATGATAGCTATTTGGGACCATGAAAGTAGTGTGTTTCTTCACTTCTGCT 1205
Db 919 CAAACAAAGAGTTATTTGGCCTTGGCTTATCAAAATATATGGGTTTCAITTTCTCTCGAT 978
Qy 1206 ACCTCACAACAGGACCATTTTCGGCTTCGGCTGTGAACATATAACGCTGATGCAAGACAG 1265
Db 979 ATCTGCTACAGGCTCTCTTTCTAGGCTGCTGTGAATCATGAAGCGGGGCAAGACTG 1038
Qy 1266 CAGCTTCCAAACATTAATATGTCATTTGACATTAATTTGACATTTGTTATTTCTGACACCT 1325
Db 1039 GATTTATCAGGAATCATATGGCATAATAATTTGACAGTGTCTCTTTGTTATGACACCAT 1098
Qy 1326 TGTTCATTAACACTCCCTGGTGGTGTATCAGCTATTTATCGTATTTCTGTAATGCTTGAC 1385
Db 1099 TATTTACTGATATACCTCAGTGTGCATTTGGCTGCCAATTTGTGATTTCTGCTGCTGAC 1158
Qy 1386 TCATAGATTTCAAGCAGCCATCCATCTATTAAAGGTTGACAAAATTTGACATTTTGGTGT 1445
Db 1159 TGGTAGATTGAAGAGGCCATCTTCCCTGGGGGTATTGATAAGAGGATTTCTTCTGT 1218
Qy 1446 GCATCAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTTAGTTCATAGCTA 1505
Db 1219 GGGCGATGACATTTACTACAACCTTAACTTTTGGCAATTTGAGATTTGTTGTTGGG 1278
Qy 1506 TTGTAATATCTGACTTCGGTACTTCTATTATTTATTTGAAGGCCAAGGACATTCGTTTGG 1565
Db 1279 TCGGGTTTTGCTGGCAATTTGTGATCCATGAATCTGCAAAATCCGCAATATAGTGTGTTGG 1338
Qy 1566 GCACATTTCCAAATTTCTGTGATATACCGAAATTTGAGCAGCTATCAAAATGCAAAACATG 1625
Db 1339 GCGGTTTTGCTGGCCACCTGTTGTATACGGAATACATTCAGTACCTCAGGCTTATACAT 1398
Qy 1626 TTCTCGAATGCTAATTTCTAGAGATTTGATGCACCAATTTACTTTTGCCAAATGCGACTAT 1685
Db 1399 ACACGGGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1458
Qy 1686 TAAGAGAAGGATCAAGAGTGGATTTGATGAAGAGAAGAAAGAAATTTAAAGCTACAGGG 1745
Db 1459 TAAAGGACAGGTTGGCTGAGTATGAGCTCAAACTCCCAATTTCAAAACCGTGGACCTGATG 1518
Qy 1746 AGACTAGTTTGCAGTATGTTAATTTGATATGAGTGTGTTGGAACATTTGATACAAGTG 1805
Db 1519 TTGAAGGAGTGTACTTTTGTGATCTCCGAGATGTCCTTGTGTTACATACATCTGAGG 1578
Qy 1806 GAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGCTACAGCTTGT 1865

Db 1579 CTGTTCAAGCTCTCAAGGACCTGCACCAAGAATACAAAGCACGACATCCAGATTGCTA 1638
Qy 1866 TGTGCAATCTGTAAGTGAAGTATGAAGAACTGAACAAATCGAAGTTCCAAAATCAT 1925
Db 1639 TAGCGAATCTTACCGGAGGTGACCTATTGCTGTCAAGAGCGGCATCATGACATGA 1698
Qy 1926 TAGGAAGAAATGGATCTATCTGACTGTTGAAGAGCGCGTTGGAGCATCAAACTTCAATC 1985
Db 1699 TTGGCGCAGGTTGTTGTTGTCGAGTGACACGCGGTGCAAGTATGCTCCAGCATG 1758
Qy 1986 TACG 1989
Db 1759 TGCG 1762

RESULT 9
US-09-720-317A-13
; Sequence 13, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-720-317A-13

Query Match 4.8%; Score 118; DB 4; Length 493;
Best Local Similarity 55.0%; Pred. No. 1.7e-22;
Matches 252; Conservative 0; Mismatches 205; Indels 1; Gaps 1;
Qy 816 TAGGATGTGTCTTCATTTCTCTCTTAGCACAAAGATACATTACGAAAAAACGACCAA 875
Db 36 TCGGATGCTCTCTCTCATATTCTCTCACCACACGGTTTCATCGGAGAGGTACAAGA 95
Qy 876 GGTGTTTTGGGTGTCAGCAATGGCGCATTTGACGTCGCTGTATATTGGGAAGTCTCTGG 935
Db 96 AGCTGTTCTGGCTGTGAGCGATCTCGCTCTGCTGTGGTCTCTCTGTCACCGCTGCGG 155
Qy 936 TTATTTTCACTACGCGGAGAGCAGGTGTTGAAGTATAGGAGACTGAAGAGGTT 995
Db 156 TCTACCGCAAGAGGCTGACAGGCAAGCGGTCAAGATCATCCAGAGGTGCAAGCGGCC 215
Qy 996 TGAATCCACCATCACT-CACAAATCTGGTATTGTGTGCTCTTACATGACTACAGCTGC 1054
Db 216 TAAACCAAGCTCCGTGGAGAGAGATACCTCAACGCGCGGACACAAAGGAGTGGCC 275
Qy 1055 AAAAAGTGGCATTTGCTGGCATCATATCACTTTCGGAAGGAATAGCAGTAGGAAGC 1114
Db 276 CAAGATCGCGGTCTATCTGCGCATCATCGCCCTCACGGAAGTATCGCGTTGGCCGATCT 335
Qy 1115 TTGCAATGTATAAAAATATACATATTGATGGCAACAAAGAGATGATAGTATTGGACC 1174
Db 336 TTGCGCTCCGTAAGAGGGTACAGACTCGACGGCAACAAAGAGATGCTGGCCATGGGGTTC 395
Qy 1175 ATCAAGTGTGTTGTTCTTTCAGCTCTTGTCTACTCACAACAGGACATTTTTCGCGTTCG 1234
Db 396 TCCAAGTGTGCTGGTCTCTGTCTCTGTGTATGTGGCAACAGGTTCTCTCCGCAAGC 455
Qy 1235 GCTGTGAATATTAACGCTGGATGCAAGACAGCAGCTTC 1272
Db 456 GCAGTGAACCTTCAGCGCGGGGCCAGTCGACCGTTTC 493

RESULT 10
US-09-720-317A-9
; Sequence 9, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (398)
; OTHER INFORMATION: n = A, C, G or T
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (483)
; OTHER INFORMATION: n = A, C, G or T
US-09-720-317A-9

Query Match 4.8%; Score 117.4; DB 4; Length 484;
Best Local Similarity 55.7%; Pred. No. 2.5e-22;
Matches 265; Conservative 0; Mismatches 207; Indels 4; Gaps 2;
Qy 717 TTCTTGGCTTGAGCATTTCCACCATGGAGCTGATATCATATCAGTGATCGCCTCTGTTT 776
Db 2 TTACAGGCATCAAGAGCTTTACAAAGAAACCGATATAATTTCCGTGATGAGCTGAAGTC 61
Qy 777 TCACCAAACTCATGA---GTGGAGGTGGGAAAGTCTGTGTAGTAGTGTCTTTCATTT 833
Db 62 CAAACAGAGCGCATAAACAGGTGGAAATGGCAAACTATTGTGATGGCATAACTTTCCCTTG 121
Qy 834 TCTTCTCTCTTAGCAAGATACCTTTCAGCAAAACGCAAGGTTTTTTTGGGTGTCTAG 893
Db 122 CATTCTCTTCTGCTTCCCAAGTACATGGAAAGAGAAATAGGAAGTTCTTCTGGTGCCAG 181
Qy 894 CAAATGGCGCATTTGACGTCCTGTTATATTGGGAAGTCTCTTGGTTTTTATTTCACTCAGCGC 953
Db 182 CTATTGCTCTTATACTTCAGTTATTTTGGCAACCTTTTGTGTTCTATTCTCGTGTG 241
Qy 954 AGAAGCAGCGTGTGGAAGTATAGAGAACTGAAGAGGTTTGAATCCACCATCACTCA 1013
Db 242 ACAAGCAAGGTGTCAGATTGTTAAACCATCAAAAGGGGCAATAAACCCATCATCAGTCC 301
Qy 1014 CAAATCTGGTATTGTGTCGCTTACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073
Db 302 ACAAAATTTATTTCATCTGCTCATTGTTGCAAAAGGTTTCAAGATCGGTGTCATTTCG 361
Qy 1074 GCATCATATCACTTCGGAAGGAATAGCAGTAGGAAGAGCTTTTGCATATGTATAAAAT 1133
Db 362 CCATGATCGTTTAAACGAGCTGTGGCAATTTGGGNGACGTTTGTCTGCTCTGAGGACT 421
Qy 1134 ACAATTTGATGGCAACAAAGAGATGATGTTTGGGACCAATGAACAGTATGTTGGT 1189
Db 422 ATCAATTTAGATTGGAACAGGAGATG-GAAACATTTGGAACATATGAACATACAAGGT 476

RESULT 11
US-09-785-381-4
; Sequence 4, Application US/09785381
; Patent No. 6602992

```
; GENERAL INFORMATION:
; APPLICANT: DALLOS, Peter
; APPLICANT: ZHENG, Jing
; APPLICANT: MADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-3701
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2441
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-785-381-4

Query Match      3.7%; Score 90.4; DB 4; Length 2441;
Best Local Similarity 46.9%; Pred. No. 2e-14;
Matches 317; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

QY 1061 GGCATTGTCGTTGGCATCATATCACATTGGGGAAGGAATAGCAGTAGGAAGAGCTTTGCA 1120
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1121 ATGTATPAAAAATTPACAATATTTGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAAC 1180
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1293 AATAAGCATGGCTACCAGGTTGATGGCAATCAGGAGCTCATTTGCCCTTGGGGATATGCAAC 1352
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1181 GTAGTTGGTCTTTACCTCTTGCTACCTCACAAAGGACCAATTTCCGGCTTGGGCTGTG 1240
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1353 TCCATTGGATCTCTCTTCCAAACCTTCTCGATTCTCTGCTCTGCTGCTCGAAGCCCTTGT 1412
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1241 AACTATAACGCTGGATGCAAGACAGCAGCTTCCAAACATTATAATGTCACCTTCGAGTAATG 1300
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1413 CAGGAAGGAACCTGGAGGGAACACAGCTTGAGGTTGTTGGCCCTCGTTGATGATCTG 1472
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1301 TTGACATTGTTAFTCCAGACCCCTTGTTCATTAACATCCCTCGTGGTGGCTATCAGCT 1360
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1473 TTGGTCATATTAGCCACCGAATTCCTCTTTGAGTGTGTACCCAGGCTGTCTTTCCGCG 1532
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1361 ATTATCGTATCGCAATGCTTGGACTCATAGATTGAAGCAGCCATCCATCTATTTAAG 1420
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1533 ATTGTGATCGTCAACCTGAAAGGAATGTTTCATGCAAGTTCTCAGACCTGCCTTTTFTTGG 1592
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1421 GTTGA---CAAAATTTCACATTTTGGTGTGCATGAGTGCATACATTGGCGTGTCTTTGGC 1477
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1593 AGNACAGCAAAATAGAGCTGACCATCTGGCTGACCACTTGTGTCTCCCTGTTCTCCTC 1652
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1478 AGTGTGAAATTGGCTTAGTCATPAGCTATTGTAATATCTGTACTTCGGGTACTTCTATT 1537
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1653 GGCTTGGACTAGCGACTGATACCGCGTGATCATTTGCTCTGCTCAGCTGATTTATAGA 1712
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1538 ATTGCAAGGCAAGGACATTCGTTTGGGCAACATTCCAAATTCGTGTATATACCGAAT 1597
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1713 ACACAGAGTCCAAAGCTACAAAGTCTCTGGGCGAGCTCCCTGACACGGATGTACATGAC 1772
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1598 GTTGAGCACTATCAAAATGCAAAACATGTTCTCGGAATGCTAATCTCAGAGATTGATGCA 1657
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1773 ATAGATGCAATAGGAGGTGAAGAAATTCCTTGGAAATTAATAATATCCAAATTAATGCC 1832
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1658 CCAATTTACTTTGCCAATGCCAGCTATTTAAGAGAAAGGATCACAAGGTGGATTGATGAA 1717
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1833 CCAATTTACTACTCAATAGCGACTTGTATAGCAGCGCTTTTAAAGAAAGAAAGCTGGAGTA 1892
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1718 GAAGAAGAAAGAAATTA 1733
Db      ||||| ||||| |||||
QY 1893 AACCCAGCACTCATTA 1908
Db      ||||| ||||| |||||

; Sequence 5627, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5627
; LENGTH: 4927
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5627

Query Match      3.4%; Score 82.6; DB 4; Length 4927;
Best Local Similarity 47.0%; Pred. No. 4.3e-12;
Matches 290; Conservative 0; Mismatches 324; Indels 3; Gaps 1;

QY 1087 TCGGGAAGGAATPAGCAGTAGGAAGAAGCTTTGCAATGTATATAAATTAACAATATTGATGG 1146
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1307 TGCTATTGCAGTGTCTAGTAGGAAAAAGTATATGCCACCAAGTATGATTACCCATCGATGG 1366
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1147 CAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTCACCTCTTGCTA 1206
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1367 GAACCAAGGAATTCATTGGCTTTGGGATCAGCAACATCTTCTCAGGATTTCTCTCTTGT 1426
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1207 CCTCAACACGAGCACTTTTCGCTTTCGCTGTGAACATATAACGCTGATGCAAGCACAGC 1266
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1427 TGTGGCCACCACTGCTCTTTCCGCAACGCGCTCCAGAGACACTGAGGAAACACACA 1486
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1267 AGCTTCCAAACATATAAATGTCACTTGCAGTAATGTTGAATTTTATTCCTGACACCTTT 1326
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1487 GGTTCCTGGCATCATCTCTGCTGCGATTGTGATGATCGCCATTTCTGCTGGGAAGCT 1546
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1327 GTTCCATTACACTCCCTCGTGGTGTCTATCAGCTATTATCGTATCTGCAATGCTTGGACT 1386
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1547 TCTGGAACCTTTGCAGAAAGTCGGTCTTGGCAGCTGTTGTAATTGCCAACCTGAAAAGGGAT 1606
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1387 CATAGATTATGAAGCAGCCAT---CCATCTATTAAAGTTGACAAATTTGACTTTTGTGT 1443
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1607 GTTTATGCAGCTGTGTGACATTCCTCGTCTGTGGAGACAGATAAGATTGATGCTGTTAT 1666
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1444 GTGCATAGAGTCATACATTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGC 1503
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1667 CTGGGTGTTTACGTGTATAGTGTCCATCATTTCTGGGCTGGATCTCGGTTTACTAGCTGG 1726
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1504 TATTGTATATCTGTACTTCTGGGTACTTCTATTATTTGCAAGGCCAAGACATTCGTTTT 1563
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1727 CCTTATATTGGACTGTGACTGTGGTCTCTGAGAGTTTCAGTTTCTTCTTGAATGGCCT 1786
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1564 GGGCAACATTCCAAATTCGTGATATACGAAATGTTGAGCACTATCAAAATGCAAAACA 1623
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1787 TGGAAAGCATCCCTAGCACAGATATCTAANAAGTACCAAGATTAACAANAACATTGAAGA 1846
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1624 TGTCTCTGGAATGCTAAATTTCTAGAGATTGATGACCAATTTTACTTTTGCAATGCCAGCTA 1683
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1847 ACCTCAAGGAGTGAAGATTCTTAGATTTCAGAGTCTTATTTTCTATGGAATGTCGATGG 1906
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

; TTTAAGAGAAAGGATCA 1700
; ||||| ||||| |||||
; TTTTAAAAAATGTATCA 1923
; ||||| ||||| |||||

RESULT 12
US-09-949-016-5627
```

US-09-785-381-2
; Sequence 2, Application US/09785381
; Patent No. 6602992
; GENERAL INFORMATION:
; APPLICANT: DALLOS, Peter
; APPLICANT: ZHENG, Jing
; APPLICANT: MADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-3701
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 4113
; TYPE: DNA
; ORGANISM: Meriones unguiculatus
US-09-785-381-2

Query Match 3.3%; Score 79.8; DB 4; Length 4113;
Best Local Similarity 46.1%; Pred. No. 2.3e-11;
Matches 304; Conservative 0; Mismatches 352; Indels 3; Gaps 1;
QY 1060 TGGCATTGTCGTTGGCATCATATCACTTGGGAAGGAATAGCAGTAGGAAGAGCTTTGCG 1119
DB 1249 TGCATTTGCCATAGCATTGTTGGATTTTCAGTGACATTTTCATGCCAACAACCTTGGC 1308
QY 1120 AATGTATAAAATTAACAATATTGATGGCAACAAGAGATGATAGCTATTTTGGGACCATGAA 1179
DB 1309 GAATAAGCATGGCTACCAAGTTGATGGCAATCAGGAGCTCATGCTTTGGGGATATGCAA 1368
QY 1180 CGTAGTTGTTCTTTTACCTCTTGCTACCTCAACAAGAGACCATTTTTCGGTTCGGCTGT 1239
DB 1369 CTCATCGGATCTCTCTCCAGACCTTCTCCATTTCTGCTCTGCTCTCGAGCCCTTGT 1428
QY 1240 GAACTATAAGCTGGATGCAAGACAGCAGCTTCCAACTATTATATGCTGCACTTGCATTAAT 1299
DB 1429 TCAGGAGGNACTGGAGGAGAAACACAGCTCGCAGGTGCTGGCCCTCGCTGATGATCT 1488
QY 1300 GTTGACATTTGTTATCTCTGACACCTTTGTTCCATTACACTCCCTCGTGGTGTCTATCAGC 1359
DB 1489 GCTGGTCACTTTAGCCACTGGATTCTCTTTGAGTCACTTGGCCCAAGGCTGTCTCGGC 1548
QY 1360 TATTATCGTATCGCAATGCTTGGACTCATAGATTATGAGCAGCCATCCATCTATTAA 1419
DB 1549 CATTGTGATCGTGAACCTGAAAGGGATGTTTATGCAAGTTCTCAGATCTGCCCTTCTCTG 1608
QY 1420 GGT---GACAAATTTGACTTTGTGTGTGATGATGATGATGATGATGATGATGATGATGAT 1476
DB 1609 GAGAACCAAAATAGAGCTGACCACTGGCTTACCACTTGTGTCTCTCCCTGTTCT 1668
QY 1477 CAGTGTGAAATGGCTTAGTCATAGCTATTGTAATATCTGATCTTCGGGTACTTCTATT 1536
DB 1669 GGGCTTGGACTACGGACTGATTTACTGCTGTGATCATTTGCTCTGCTGACTGTGATTACAG 1728
QY 1537 TATTGCAAGCCCAAGGACATTCGTTTGGGCAACATTCCAAATCTGTGATATACCGAAA 1596
DB 1729 AACCCAGAGTCCAGCTACAAAGTCTCGGGGAGCTCCCTGACACCGCATATACATTGA 1788
QY 1597 TGTGTAGCACTCAAAATGCAAAACATGTTCTGGAATGCTAATTTCTAGAGATTGATGC 1656
DB 1789 CATAGACGATATAGAGAGGTGAAGAAATTTCTGGAATAAAATATTCAGATAAAGC 1848
QY 1657 ACCAAATTTACTTTGCCAATGCCAGCTATTTAAGAGAAAGGATCACAAGGTGATGATG 1715
DB 1849 CCCAATTTACTATGCAACAGTCACTTGTATAGCAAGCCCTAAAGAAAGAGCTGGTG 1907

RESULT 14
US-09-248-796A-6643
; Sequence 6643, Application US/09248796A

; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6643
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6643

Query Match 3.2%; Score 77.6; DB 4; Length 2487;
Best Local Similarity 48.8%; Pred. No. 7.2e-11;
Matches 209; Conservative 0; Mismatches 219; Indels 0; Gaps 0;
QY 257 CTTCAAGTTCTTCTCCCATTTTCGAATGGGCTCCCAATACACCTTTTCAGTTCTTGAAA 316
DB 253 CTTAATTTCAATGTTTCCCAATTTCTTAATGGAATTTTGCATTAACAATATTAGATGGTTATAT 312
QY 317 GCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAGGGCATCAGTTAT 376
DB 313 GGTGACTTGGTTGGTGAATTAATCTGTTGGAGTTGTTTGTAGTGCCTCAATCTATGTCATAT 372
QY 377 GCCAAGCTGCCAACCTCCCTCCAATTTCTTGACTATATTCGAGCTTTTATACCAACCATG 436
DB 373 GCCCAATTAGCTGGGTAGAGCTCAATATGTTTATATCTCTCTTTGTCGGTGTTC 432
QY 437 ATTTATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTGTGGCGGTGGATCG 496
DB 433 ATTTATTCATTTTTCGCCACCTCAAAAGATGTTTCCATCGGCTCTGCTGTATGTCC 492
QY 497 CTTCTGATGGGTTCATGTTGATGTAATGCCGTTGATCCCATGAAGACCCAAAGCTTTAC 556
DB 493 TTGCAAGTGTCTAAGGTCAATGCTCATGTTCAAGATAAGTTTGGTGACAAGTATGCTGCT 552
QY 557 CTCACCTGGCTTTTCACAGCTACATTTATTTGCTGGTGTGTTTTTCAGGCTGCCCTTGGGCTG 616
DB 553 CCTGAAATGCGCAATTTTGTCAATGATTTGTTGGTGGTATCGCTCTTGGTATGTTGTTG 612
QY 617 TTTAGTTGGGGTTGATCGTGATTTTCTGTCATGCAACCATTAATAGGGTTTCATGGGA 676
DB 613 TTACGTTTAGGATTCATTTTAGAATTTATTTCTATCCAGCGGTCAATGGGGTTCAATGACT 672
QY 677 GGAGCAGC 684
DB 673 GGTCTGTC 680

RESULT 15
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 11:07:04 ; Search time 5638 Seconds
(without alignments)
2852.638 Million cell updates/sec

Title: US-10-762-049-17

Perfect score: 2449

Sequence: 1 gcaagcgtagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1521.6	62.1	2008	18	US-10-424-599-131116 Sequence 131116.
2	1018.2	41.6	1977	9	US-09-938-842A-1502 Sequence 1502, Ap
3	1018.2	41.6	1977	11	US-09-938-842A-1502 Sequence 1502, Ap
4	846	34.5	2372	19	US-10-437-963-38184 Sequence 38184, A
5	790.2	32.3	1518	18	US-10-424-599-131117 Sequence 131117,
6	677.4	27.7	2727	20	US-10-425-115-50601 Sequence 50601, A
7	670.4	27.4	4330	19	US-10-437-963-77012 Sequence 77012, A

8	658.6	26.9	2366	18	US-10-424-599-65103 Sequence 65103, A
9	650.6	26.6	2651	20	US-10-425-115-182207 Sequence 182207,
10	647.2	26.4	2780	19	US-10-437-963-77005 Sequence 77005, A
11	600.6	24.5	2001	19	US-10-437-963-95853 Sequence 95853, A
12	571.4	23.3	2429	19	US-10-437-963-73410 Sequence 73410, A
13	567	23.2	2856	18	US-10-425-114-15719 Sequence 15719, A
14	567	23.2	2713	20	US-10-425-115-75763 Sequence 75763, A
15	520.2	21.2	536	18	US-10-424-599-32308 Sequence 32308, A
16	499.6	20.4	680	18	US-10-424-599-32307 Sequence 32307, A
17	489	20.0	2336	18	US-10-424-599-52175 Sequence 52175, A
18	482.2	19.7	3201	20	US-10-425-115-44188 Sequence 44188, A
19	477.4	19.5	1992	18	US-10-425-114-13333 Sequence 13333, A
20	465.2	19.0	2242	20	US-10-425-115-40196 Sequence 40196, A
21	433.8	17.7	585	18	US-10-424-599-108738 Sequence 108738,
22	414.6	16.9	2251	19	US-10-437-963-72062 Sequence 72062, A
23	401.4	16.4	1980	17	US-10-260-238-1066 Sequence 1066, Ap
24	396.8	16.2	1809	18	US-10-424-599-83137 Sequence 83137, A
25	389.6	15.9	2506	18	US-10-424-599-74393 Sequence 74393, A
26	372.4	15.2	1303	18	US-10-424-599-139734 Sequence 139734,
27	361.8	14.8	1247	18	US-10-424-599-130951 Sequence 130951,
28	361.2	14.7	2156	18	US-10-424-599-60281 Sequence 60281, A
29	359.6	14.7	2156	20	US-10-739-930-3837 Sequence 3837, Ap
30	348.8	14.2	1551	18	US-10-424-599-54350 Sequence 54350, A
31	339.8	13.9	674	19	US-10-767-701-9435 Sequence 701, App
32	329.4	13.5	1115	19	US-10-767-701-9435 Sequence 9435, Ap
33	323.8	13.2	3252	19	US-10-437-963-78357 Sequence 78357, A
34	315.4	12.9	1019	19	US-10-767-701-11661 Sequence 11661, A
35	311.6	12.7	1069	20	US-10-425-115-3619 Sequence 3619, Ap
36	298.6	12.2	2913	19	US-10-437-963-76453 Sequence 76453, A
37	298.2	12.2	1174	20	US-10-425-115-75765 Sequence 75765, A
38	293.2	12.0	2346	20	US-10-425-115-22539 Sequence 22539, A
39	291.4	11.9	1632	18	US-10-424-599-74389 Sequence 74389, A
40	285.6	11.7	565	18	US-10-424-599-62889 Sequence 62889, A
41	271.4	11.1	788	20	US-10-425-115-75767 Sequence 75767, A
42	258.8	10.6	1883	18	US-10-424-599-103813 Sequence 103813,
43	255	10.4	459	18	US-10-424-599-120407 Sequence 120407,
44	249.8	10.2	735	18	US-10-424-599-35721 Sequence 35721, A
45	249.4	10.2	841	18	US-10-424-599-135930 Sequence 135930,

ALIGNMENTS

RESULT 1

US-10-424-599-131116
; Sequence 131116, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131116
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89404C.1
US-10-424-599-131116

Query Match 62.1%; Score 1521.6; DB 18; Length 2008;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 84; Indels 2; Gaps 1;
QY 780 CCCAACTCATGAGTGGAGGCGGAAAGTCGTGTGTAGATGTGCTTCATTTCTTCC 839
|||||
Db 1 CCCAAAGCATGAGTGGAGGCTGTGAAAGCGCTGTGTGGTGTGCTGATTTCTTCC 60
|||||

Qy	840	TCCTTAGCACAAGATACCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTCAGCAATGG	899
Db	61	TCCTCGTCAAGATACCTTCAGCAAAACGACAGCAAAAGTTCTTCTGGGTCTCAGCAATGG	120
Qy	900	CGCAATTGACGTCGGTTATATTTGGGAAGTCTCTTTGGTTATTTTCACTCACGCCGAGAAAGC	959
Db	121	CACCACATGACGTCGGTTATATTTGGGAAGTCTGCTGGTCTATGTGCACACACGCTGAGAAAC	180
Qy	960	ACGGTGTGAA--GTGATAGGAGAACTGAAGAGGGTTGAAATCCACCACATCACTCAAAA	1017
Db	181	ATGAGATTCAGGTGTGATAGAGAAATTTGAAGAGGGTTTGAATCCACCATCACTGACAAA	240
Qy	1018	TCTGGTATTTGTCGTGGCCTTACATGACTACAGCTGTCAAAACTGGCAATTTGTCGTGGCAT	1077
Db	241	TCTGGTATTTGTGACGCTTACATGACCAACAGCTGTCAAAACAGGCATTTGTCGTGGCAT	300
Qy	1078	CATATCACTTCGCGAAGGAATAGCAGTAGGAGAAAGCTTTTGCAATGTATAAAAATTACAA	1137
Db	301	TATATCGCTTCGCGAAGGAATAGCAGTGGGAAGAGCTTTTGCAATGTATAAAAATTACAA	360
Qy	1138	TATTGATGGCAACAAAGAGATGATAGCTATTTGGGACCATGAAACGTAGTCTGTTCTTTCCAC	1197
Db	361	TATTGATGGCAACAAAGAGATGATAGCTATTTGGGACCATGAAACGTAGTCTGTTCTTTCCAC	420
Qy	1198	CTCTTGCTACCTCAACACAGSACCATTTCGCGTTCCGGTGTGAACTATTAACCGCTGGATG	1257
Db	421	CTCTTGCTACCTCACAACTGACCATTTCGCGTTTCAGCTGTGAACTATAACCGCTGGATG	480
Qy	1258	CAAGACAGCAGCTTCCAACATTAAATGTCTCACTTGCAGTAATGTTGACATGTTATTTCCCT	1317
Db	481	CAAGACAGCAGCATCCAACATCAATATGTCATTTGCAAGTAAATGTTGACATGTTATTTCCCT	540
Qy	1318	GACACCTCTGTTCCATTACACTCCCTCGGTGGTCTATCAGCTATTAATCGTATCTGCAAT	1377
Db	541	AACACCAATGTTCCATTACACTCCCTCGGTGGTCTATCAGCCATTATTGTAICTGCAAT	600
Qy	1378	GCTTGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGTTGACAAAATTTGACTT	1437
Db	601	GCTTGGCTCATAGATTATGAAGCAGCCATCCATCTATTTAAGTTCGACAAAATTTGACTT	660
Qy	1438	TGTGGTCTGCATGAGTCGCATACATTGGCGTGTCTTTGGCAGTGTTCGAAATTCGCTTAGT	1497
Db	661	TGTGGTCTGCATGAGTCGCATGACGTGGCGTGTCTTTGGCAGTGTTCGAAATTTGCTTAGT	720
Qy	1498	CATAGCTATTGTAATATCTGTACTTCGGGTACTTCTATTTATTCGAAGGCCAAGGACATT	1557
Db	721	CATAGCTATTGCGATATCTGTACTTCGGGTACTTCTATTTATTCGAAGGCCAAGGACATT	780
Qy	1558	CGTTTTGGGCAACATTCGAAATTCGTGTGATATACCGAAATGTTGAGCAGCTATCAAAATGC	1617
Db	781	TGTTTTGGGAAACATTCGAAATTCGTGTGATATACAGAAATGTTGAGCAGCTATCCAAATGC	840
Qy	1618	AAAACATGTTCTCGAATGCTAAATTCATAGAGATTGATGCACCAATTTTACTTTTGCCAAATGC	1677
Db	841	AAAACATGTTCTCGAATGCTAAATTCATAGAGATTGATGCACCAATTTTACTTTTGCCAAATGC	900
Qy	1678	CAGCTATTTAAGAGAAAAGGATCAACAAGTGGATTGATGAAGAGAGAAAGAAATTTAAAGC	1737
Db	901	AAGCTATTTAAGAGAAAAGGATCAACAAGTGGATTGATGAAGAGAGAAAGAAATTTAAAGC	960
Qy	1738	TACAGGGGAGACTAGTTTGCAGTATGTTATAATTTGATATGAGTCTCTTTCGGAACATTGA	1797
Db	961	TACAGGGGAAACTAGTTTGCAGTATGTTATAATTTGATATGAGTCTCTTTCGTAACATTGA	1020
Qy	1798	TACAAGTGGAAATAAGTATGCTTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGCTACA	1857
Db	1021	TACAAGTGGAAATAAGTATGCTTTGAAGAGGTGAAGAAGATTATAGAGAGAGAGAGTTACA	1080
Qy	1858	GCCTGTTTTGGTCAATCTCTGTAAGTGAAGTGATGAAGAACTGAACAAATTCGAAGTTCCA	1917
Db	1081	ACTTGTGTTTTGGTCAATCTCTGTAAGTGAAGTGATGAAGAACTGAACAAATTCGAAGTTCCA	1140

Qy	1918	AAATCATTTTAGGGAAGAAATGGAATCTATCTGA	CTGTTGAAGAGCGCTTTGGAGCATGCAA	1977
Db	1141	AAATCATTTTAGGGAAGAAATGGAATCTATCTGA	CTGTTGAAGAGCGCTTTGGAGCATGCAA	1200
Qy	1978	CTTCAATCTACGTGCAAGCAAAACGAAACCAA	AGAAAGATGAAACAGAGGGTTGGAAACAA	2037
Db	1201	CTTCAATCTACGTGCAAGCAAAACGAAACCAA	AGAAAGATGAAACAGAGGGTTGGAAACAA	1260
Qy	2038	TGTGTGACTCAGTCATATGCCAAAGAGTATTC	TAAATAAATCTCAAAAAGCTTATTCGTTTT	2097
Db	1261	TGTGTGACTCAGTCATATGCCAAAGAGTATTC	TAAATAAATCTCAAAAAGCTTATTCGTTTT	1320
Qy	2098	CGCTTAGTAAATGTTTACCATCTACAATGTG	GCGCATGAGAAATTTCTGAATCACGCCGAAGA	2157
Db	1321	CGCTTAGTAAATGTTTACCATCTACAATGTG	GCGCATGAGAAATTTCTGAATCACGCCGAAGA	1380
Qy	2158	AGTTTTAAAGGCATAGGAAAAATGAAAGATG	CAAGGGCTTCTTAATTTCTCAACTCTGCGAT	2217
Db	1381	AGTTTTAAAGGCATAGGAAAAATGAAAGATG	CAAGGGCTTCTTAATTTCTCAACTCTGCGAT	1440
Qy	2218	CCTTAGTTTGAAGAAATCTCCATGTGTATAGG	CTGTTGAAATAATCTTTTACGTATCATGCG	2277
Db	1441	CCTTAGTTTGAAGAAATCTCCATGTGTATAGG	CTGTTGAAATAATCTTTTACGTATCATGCG	1500
Qy	2278	TTGATAATATATTCAAGAGAGAAATGCTAG	CAACACACTCTCAGACACACTCTTTTGAACAC	2337
Db	1501	TTGATAATATATTCAAGAGAGAAATGCTAG	CAACACACTCTCAGACACACTCTTTTGAACAC	1560
Qy	2338	ATGTAAAGAGGTAAAGAAAGTGTGTTGCTAG	CACCTCCATATTTCAATTTGAAGTAAATT	2397
Db	1561	ATGTAAAGAGGTAAAGAAAGTGTGTTGCTAG	CACCTCCATATTTCAATTTGAAGTAAATT	1620
Qy	2398	GCCATGAGAAATTTAAAAATCTCTTTGGAAAA	AAAAAAAAAAAAAAAAAAAAA	2447
Db	1621	GCCATGAGAAATTTAAAAATCTCTTTGGAAAA	AAAAAAAAAAAAAAAAAAAAA	1670

RESULT 2

US-09-938-842A-1502
 ; Sequence 1502, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRI1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1502
 ; LENGTH: 1977
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1502

	Query Match	41.6%;	Score 1018.2;	DB 9;	Length 1977;
	Best Local Similarity	72.3%;	Pred. No. 4.1e-263;		
	Matches 1364;	Conservative	0;	Mismatches 518;	Indels 5; Gaps 3;
125	CACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTCAAG				184
58	CACACGGTGGAGGCTCCACAACCTCAACCGTTCTTTGAAGTCACATTCAGTACTCAGTGAAG				117
185	GAGACTTTCTTCCCTGGATGACCCCTTTGAGGCAGGTTTCAAGAACACAGCCAGCGTTTCCAAGAAG				244

Db 118 GAAACTCTGTTCCAGACGACCCCTTTTAGACAATTTAAGAACCAAAAATGCATCAAGAAA 177
Qy 245 TTCATGCTGGCCCTTCAGTCTCTCTCCCATTTTCGAATGGGCTCCCAATACACCTTT 304
Db 178 TTGTGTAGGCTCAAAATCTCTCTCCGCAATTTTGAATGGGCAACAGCTACAATCTC 237
Qy 305 CAGTCTCTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAG 364
Db 238 AAGTCTCTCAATCAGATCTCATCGCGGAATCACCATCGTAGTCTCGCCATCCCTCAG 297
Qy 365 GGATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTCTGGACATATATTCGAGCTTT 424
Db 298 GGATCAGTTATGCCAACCTTGTAACTTGTCCGCCCAATCTCTGGCCCTTTATTCGAGTTT 357
Qy 425 ATACCAACCATGATTTATCGGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAGCTGG 484
Db 358 GTACCCGCTATGATACGCGGTGCTAGGAGTTCAAGGGACTTAGCGGTGGGAACGGTT 417
Qy 485 GCGGTTGGATCGCTTCTGATGGGTTCCGATGTTGAGTAATGCCGTTGATCCCAATGAAGAC 544
Db 418 GCGGTTGCGTCTCTGTTGACAGGTGCGATGCTGAGCAAGAAGTTGATGCTGAGAAGAT 477
Qy 545 CCAAGCTTTACCTCCAGCTGGCTTTCAGAGTACATATTTGCTGGTGT-TTTTGAGGC 603
Db 478 CTTAAGCTTTACCTTACCTTGTCTTTCAGCGGCACCTTTTTCGCGCGCTTCTCGAAGCC 537
Qy 604 TGCTTTGGGCTCTGTTAGGTTGGGTTGATCGTGGATTTTCTGTCATGCAACCATAT 563
Db 538 TCTCTTGGAAATTTTCAAGGGTTAGGTTTCATAGTGGACTTTCTATCGCATGCAACGATGT 597
Qy 664 AGGTTTCATGGGAGGACGACCAAGCTGGTGTCTGTCAGCAACTAAATCGATTTCTGG 723
Db 598 AGAATTCATGGGAGGAGCAGCGTGGTGTGAGTCTGCAACAGCTTAAAGGTTATTTTCG 657
Qy 724 CTTTGAGCTTTTACCCATGGAGCTCATATCATATCAGTGATGCGCTCTGTTTTCACCCA 783
Db 658 ACTTAAACATTTTACAGACTTACCGATGTTATCTCTGTCATGCGTTCGGTTTCTCCCA 717
Qy 784 AACTCATGAGTGGAGTGGGAAGTCTGTTAGGATGTCATTTCTTCTCTCTCT 843
Db 718 AACTCAGAGTGAGATGGGAAGTGGCGTTCTTGGATGGGTTTCTATTTCTTCTCT 777
Qy 844 TAGCACAAGATCTTCAGCAAAAAACGACCAAGGTTTTTGGGTGTCAGCAATGGCGCC 903
Db 778 CTCCACAGATTTTACAGATCAAGAAACCAAAATTTCTTTTGGGTGGCGGATGCTCC 837
Qy 904 ATTGAGCTCGGTTATATGGGAAGTCTCTTGGTTTATTTTCACTCACGCCGAGAACGG 963
Db 838 TTTGACCTCAGTGATTTCTTGGAGTCTCTTGGTTTACTTCACTCACGCTGAGAGACATGG 897
Qy 964 TGTGAAGTATAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATCTGT 1023
Db 898 TGTTCAGTATAGGGGACCTGAAGAAAGGTTTGAATCCACTCTCCGGTTCTGATCTCAT 957
Qy 1024 ATTTGTGTCGCTTACATGACTACAGCTGTCAAAAGTGGCATGTCGTTGGCATCATATC 1083
Db 958 CTTTACTTCCCCTTACATGTCACAGCTGTCAAACTGGCCTCATCTGGCATCATTCG 1017
Qy 1084 ACTTGGGAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATFAAAAAATACAAATTTGA 1143
Db 1018 TCTCGC-GAAGGATAGCAGTGGGAGGAGTTTGGATGTTTCAAGAACTACAAACATAGA 1076
Qy 1144 TGCAACAAAGAGATATAGCTATTGGGACCATGAACGATAGTTGGTTCTTTCACCTCTTG 1203
Db 1077 CGGGAACAAAGAGATATAGCTGTTTGGAAATGATGAACATCGTTGGTTCTTTCACATCTTG 1136
Qy 1204 CTACCTCACAACAGGACCAATTTTCGCTTCCGCTGTGAACTATAACGCTGGATGCAAGAC 1263
Db 1137 TTACCTCACAACCGGACCAATTTTCAAGGTGGCAGTGAACTACACGCGGTTGCAAGAC 1196
Qy 1264 AGCAGCTTCCAAATTAATGTCACTTGCAGTAAATGTTGACATTTGTTTCTGACACC 1323

Db 1197 CGCAATGTCCAAACATAGTAGGCGATGCGGTTATGTTTCAACACTCTCTTCTCTCACACC 1256
Qy 1324 CTTGTTCCCAATTACACTCCCTCGTGGTGTATCAGCTATTATCGTATCTGCAATGCTTGG 1383
Db 1257 GCTTTTCTACTACACCACTCGTGTCTCTCTGCGCATCATATCCGCAATGCTCGG 1316
Qy 1384 ACTCATAGATATGAAGCAGCCATCCATCTATTAAAGGTTGACAAATTTGACTTTGTGT 1443
Db 1317 ACTCATGTGATATCAAGCTGCCATCCATCTCTGGAAGATTGACAAGTTCCGACTTCTCTCGT 1376
Qy 1444 GTGCGATGATGTCATCATTTGGCGTGGTCTTTGGCAGTGTGGAATTTGGCTTATCATAGC 1503
Db 1377 GTGCGATGAGCGCTACGTTGGGGTGGTATTCGGCAGTGTAGAGATTTGGACTCGTCTGAGC 1436
Qy 1504 TATTGTAAATATCTGTACTTCCGGTACTTCTATTATTTGCAAGGCCAAGACATTCGTTT 1563
Db 1437 GGTGGGATATCTATAGCGAGTTTGTGCTGTTGTGTCGAGGCCAANAACCTGCGGTGA 1496
Qy 1564 GGGCAACATTTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACA 1623
Db 1497 GGGAAACATATCCAAACAGCATGATCTATAGGAACACTGAGCAGTAGTACCCATCATCAAGAAC 1556
Qy 1624 TGTTCTCGAATGCTAAATTTCTAGAGATTTGATGCCAATTTACTTTGCCAATGCCAGCTA 1683
Db 1557 CGTTCTCGTATTTCTCATCTTGGGATTTGATGCTCCCATCTACTTTGCTTAATGCCAGTTA 1616
Qy 1684 TTTAAGAGAAAGGATCACAAGGTGATTCGATGAAGAAGAAAGAAATTTAAAGCTACAGG 1743
Db 1617 CTTGCGTGAAGAATCATTAAGGTGATTTGATGAGGAGGAGAGAGGTTTAAACATCAGG 1676
Qy 1744 GGAGACTAGTTTTCGAGTATGTTTATAATTTGATATGATGCTGTTGGAAACATTTGATACAAAG 1803
Db 1677 AGAGAGCAGCTTACAATATATTATCTCGATATGTCAGCTGTTGGTAAATATCGACACAAG 1736
Qy 1804 TCGAATAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGAGCTACAGCTTGT 1863
Db 1737 CGGTATTAGCATGATGGTGGAAATTTAAGAAAGTCTATTGACAGGAGAGCGTTAAAGTTGT 1796
Qy 1864 TTTGGTCAATCTCTGAAGTGAAGTATGAAGAACTGAACAAATCGAAGT---TCCAAA 1920
Db 1797 ATTGCAATCCAAAGGAGAGGTGCTGAGAGAAATTTACCAGATCCAAATTCATCGGTGA 1856
Qy 1921 TCATTTAGGAAAGAAATGATCTATCTGACTCTTGAAGAGCGGTTGGAGCATGCAACTTT 1980
Db 1857 TCATTTGGSCAAAGAGTGGATGTTCTTACCGTAGGAGAGCAGTGGAGGCTTGTAGCTA 1916
Qy 1981 CAATCTAGTGAAGCAAAACGAAACC 2007
Db 1917 CATGCTTACACGTTTAAACCCGAACC 1943

RESULT 3

US-09-938-842A-1502
; Sequence 1502, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1502

; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 38184
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41845C.1
US-10-437-963-38184

Query Match 34.5%; Score 846; DB 19; Length 2372;
Best Local Similarity 66.4%; Pred. No. 1.2e-216;
Matches 1231; Conservative 0; Mismatches 620; Indels 3; Gaps 1;

QY 131 GTCAGGTTCCACGCCACAGCGCTTTTCAAGTCTCTAAAGTACTTTTGAAGGAGACT 190
DB 152 GTCCGATGCCGCGGCGAAGCGCTTCTGGAGACGCTGGGGGGAACATGAAGGAGACA 211
QY 191 TTCTTCCCTGATGACCTTTGAGCGAGTTCAGAAACAGCCAGCTTCCAAAGAGTTC--- 247
DB 212 TTCTGCGGACACCGCTTTCAGGGTGGTTCGCGGCGGAGCGGGTTCGCGGCGCGCG 271
QY 248 ATGCTTGGCTTCAGTTCCTTCTCCCATTTTCGAATGGGCTCCCAAATACACTTTTCAG 307
DB 272 GCGGCGGCTCGGTACGTTCCCGTTCATGAGTGGGCGCGCTGTCACACCTTCGGC 331
QY 308 TTCTTGAAGTCTGACCTCATAGTGGGCATCACCATCGCTAGTGTGGCCATTTCTCAGGGC 367
DB 332 ACCCTCAAGTCCGACCTCATCGCGGCATCACCAATGGCAGCTCGCCATCCCGCAGGGC 391
QY 368 ATCAGTTATGCCAAGTCCGCAACCTCCCTCCAAATCTTGGACTATATTCGAGCTTATA 427
DB 392 ATCAGTACGCCAAGTCCGCAACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCG 451
QY 428 CCACCATTTGATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGAGCTGGGG 487
DB 452 CCGCGCTGTGTACCGGATGATGGGAGCTCGAGGAGCTGCGCGTGGGAGCGTGGCG 511
QY 488 GTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTTGATCCCAATGAAGACCCA 547
DB 512 GTGGCGTCTGCTGATCGGCTCGATGCTGAGCGAGGAGGTGTGCGGCGGCGGAGGCCG 571
QY 548 AAGCTTTACCTCACCTGGCTTTTCACAGCTACATTAATTTGCTGGTGTGTTTTCAGGCTGCC 607
DB 572 GCGCTGTACCTGACGCTGCGGCTCACCGCACCTTCTTCGCGCGGCTGTTCAGGCGCTG 631
QY 608 TTGGGCTGTGTTAGGTTGGGTTGATCGTGGATTTTCTGTACATGCAACCAATAATAGG 667
DB 632 CTGGGGTCTTGAGGCTGGGTTTCATCGTGGACTTCTGTGCGACGCCACCATCTGTCGGG 691
QY 668 TTGATGGGAGGAGCAGCGGTGTGTGCTGAGCAACTAAATTCGATTTCTTGGGCTTT 727
DB 692 TTGATGGGCGCGCCCGGCTGTGCTTCGAGAGTTCGAGAGTTCGAGAGTTCGAGAGTTC 751
QY 728 GAGCATTTACCATGAGCTGATATCATATCATGATGAGTGGCTCTGTTTTCACCCAAACT 787
DB 752 GACCACTTACCAACCGCCACCTCGTCTCGGTTCATGAGCTTCGCTTCTCCCAAACT 811
QY 788 CATGAGTGGAGTGGGAGTGTGTTAGGATGTCTTCATTTCTTCTTCTTCTTCTTCTTCT 847
DB 812 CACCTTGGGATGGGAGAGCTGTCATGGGCTGCGGCTTCTTCTTCTTCTTCTTCTTCTTCT 871
QY 848 ACAAGATATTTACAGCAAAACACGACCAAGGTTTTTTTGGTGTGAGCAATGGCGCCATTG 907

DB 872 ACCCGCTTCTTCAGCAAGAGGAGGCCAAGGTTCTTCTGGGTATCTGAGCTGCGCCATTG 931
QY 908 AGTCCGTTATATTGGGAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGACGCGTGT 967
DB 932 GCGTCTGTATCATGTCGGGAGCCCTCTGGTGTACCTCACATGCTGAAAACCATGCAATT 991
QY 968 GAAGTGATAGGAACTGAAGAAGGGTTTGAATTCACCATCACTCAAAATCTGGTATTT 1027
DB 992 CAAGTGATTGGTTACCTGAAGAAGGGCTGAAACCCACCGCTGCGCACAGCCCTGAACCTC 1051
QY 1028 GTGTGCGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
DB 1052 TCGTCCCGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1111
QY 1088 GCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGATATAAAATTTACAATATTTGATGGC 1147
DB 1112 GCTGAGGGAATTCGCGTAGGAAGAGCTTTGCAATGATATAAAATTTACAATATTTGATGGC 1171
QY 1148 AACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207
DB 1172 AACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
QY 1208 CTCACAAAGAGACCAATTTTCGCGTTCGCGTGTGAACATAAAGCTGGATGCAAGACAGCA 1267
DB 1232 CTCACCAAGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCG 1291
QY 1268 GCTTCCCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1327
DB 1292 ATGTGGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1351
QY 1328 TTCCATTTACCTCCCTGCTGCTATCAGCTATTTATGATGATGATGATGATGATGATGATGAT 1387
DB 1352 TTCCATTTACCTCCCTGCTGCTATCAGCTATTTATGATGATGATGATGATGATGATGATGAT 1411
QY 1388 ATAGATTTAAGAGAGCCATTCATTTTAAAGTTTGAACAATTTGACATTTTGGTGTGTC 1447
DB 1412 ATCGACTACCGCGCGCGCTTCACCTGCGAGGTGGACAAGTGGACTTCTGCGCTGTC 1471
QY 1448 ATGAGTGATACATTTGGCGTGTCTTTCGAGTGTGTAATTTGGCTTTAGTCAATGATTT 1507
DB 1472 CTCGGCGCTTACCTCGCGCTGCTTTCGCGAGCTGAGATCGCGCTGCTGCTGCGCGCTC 1531
QY 1508 GTAATATCTGTATCTCGGCTACTTCTATTTATTTGAAGGCCAAGGACATTTCTGTTTTGGG 1567
DB 1532 GGGATCTCCATCTCCGCGTGTGCTGCTGCTGCGCGCGCGCGAGGAGCGGCTGCTGCGG 1591
QY 1568 AACATTTCAAATTTCTGTGATATACCGAATGTTGAGCACTATCAAAATGCAAAACATGTT 1627
DB 1592 AACATTTCAAATTTCTGTGATATACCGAATGTTGAGCACTATCAAAATGCAAAACATGTT 1651
QY 1628 CCTGGAATGCTAATTTCTAGAGATTTGATGACCAATTTACTTTGCCAATGCCAGCTATTTA 1687
DB 1652 CCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
QY 1688 AGAGAAAGGATCACAAAGTGGATTTGATGAAGAAGAAAGAAATTTAAAGCTACAGGGGAG 1747
DB 1712 CGTGAGAGGATCGCGGCTGGATCGACGAGAGGAGGACAGTGCAGAGAGAGGAGGCGGAG 1771
QY 1748 ACTGTTTTCAGTATGTTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1807
DB 1772 ATGGGATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831
QY 1808 ATAGTATGCTTTGAAGAGGTGAAGAGATTTACAGAGAGAGAGAGCTACAGCTTGTGTTTG 1867
DB 1832 ACGAGATGCTGGATGAATCAGGAAGACCTTGGAGAGAGGCGCTTCCAGATTTGTTG 1891
QY 1868 GTCAATCTCTGTAAGTGAATGATGAAGAACTGAACAAATCTGAAGTTTCCAAATCATTTA 1927
DB 1892 GCGAACCGGAGGAGGAGATCATGAAGAGTTGGACAGCTCCAGGTGCTTGGGCGATC 1951
QY 1928 GGGAGAAATGATCTATCTGATCTGTTGAAGAGCGGCTTTGGAGCATGCAACTTC 1981
DB 1952 GGCCATGATGATCTTCCCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2005

RESULT 5
US-10-424-599-13117
; Sequence 131117, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yihwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 13117
; LENGTH: 1518
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT3847_89405C.1
US-10-424-599-13117

Query Match 32.3%; Score 790.2; DB 18; Length 1518;
Best Local Similarity 79.0%; Pred. No. 1e-201;
Matches 1033; Conservative 0; Mismatches 178; Indels 96; Gaps 4;

QY	746	GCTGATATCATATCAGTGATGCGCTCTGTTTCCACCCAACTCATGAGTGGAGTGGGAA	805
DB	24	GCTAAATAGTTTACTTGAGACATGTTTTTCCCAACGCATGAGTGGAGTGGGAA	83
QY	806	AGTGTCTGTAGGATGTCTTCAATTTCTTCCCTTAGCACAGATACCTTCAGCAA	865
DB	84	AGCGCTGTGTGGTGTGCTTATTTCTTCTCTCTCGTCAAGATACCTTCAGCAA	143
QY	866	AAACGACCAAGTTTTTGGGTGTCAGCAATGGCGCCATTCACGTCGGTATATGGGA	925
DB	144	CGACAGCAAGTTCTTCTGGGTGTGAGCAATGGCACCACTGACGTCGGTATATGGGA	203
QY	926	AGTCTCTGTTTATTTTCACTCACGCGGAGAGCAGGTGTGGAAGTATAGGAGAACTG	985
DB	204	AGTCTGCTGTCTATCTCACACAGCTGAGAAACATGGAGTTCAAGTATAGGAAATTTG	263
QY	986	AGAAGGGTTTGAATCCACATCACTCAAAATCTGGTATTTGTGTCGCTTACATGACT	1045
DB	264	AGAAGGGGTGAATCCACCTTCGGCAACGGATTTGGTGTGTGTGTACCTTACATGGGC	323
QY	1046	ACAGCTGTCAAACTGGCATTGCTGGTGCATCATATCACTTGGGAGGAATAGCAGTA	1105
DB	324	ACTGCCATAAACTGGATTCGTCATGGCAATATAGCACTTTCGGGAAGGAATAGCAGTG	383
QY	1106	GGAGAAGCTTTTCAATGTATAAAATTAATAATTTAGTGGCAACAAAGAGATGATAGCT	1165
DB	384	GGAGAAGCTTTTCAATGTATAAAATTAATAATTTAGTGGCAACAAAGAGATGATAGCT	443
QY	1166	ATTGGACCATGAACCTAGTGTGTTCTTTACCTCTTGTACTCTCAACAGACCAATTT	1225
DB	444	ATTGGAAACCATGAATATTTTGGGATCGTTTCACTTCTTGTCTATCTCACCACTGGACCAATTT	503
QY	1226	TCCGCTTCGGCTGTGAATATACGCTGGATGCAAGACAGCAGCTTCCCAACATTATAATG	1285
DB	504	TCCGCTTCAGCATGAACATACAAATGCTGGGTGCAAGACAGCTGCATCCCAACATCATATG	563
QY	1286	TCACTTGCAGTAATGTTGACATTTTATTCCTGACACCTTGTTCATTTACATCCCTCGT	1345
DB	564	GCAATTCAGTATGTTGACATTTGTTATTCCTTAACACCATTTGTTTCACTTACCTCCCTG	623
QY	1346	GTGGTGTATCAGCTATTCGTAATCTGCAATCTGGATCTGATAGATTAAGAGCAGCC	1405
DB	624	GTGGTGTATCAGCCATTTATTTGATCTGCAATGCTGGGCTCATAGATTTATGAAGCAGCC	683

QY	1406	ATCCATCTATTTAAGGTTGACAAATTTGACATTTTGGTGTGTCATGAGTGCATACATTTGGC	1465
DB	684	ATCCATCTATTTAAGGTCGACAAATTTGACATTTTGGTGTGTCATGAGTGCATACATTTGGC	743
QY	1466	GTGGTCTTTTGGCAGTGTGAAATTTGGCTTAGTATAGCTATTTGTAATATCTGTACTTGG	1525
DB	744	GTGGTCTTTTGGCAGTGTGAAATTTGGCTTAGTATAGCTATTTGCGATATCTGTACTTGG	803
QY	1526	GTACTTCTATTTATTTGCAAGGCGAAGGACATTCGTTTTTGGGCAACATTTCCAAATTTCTGT	1585
DB	804	GTACTTCTATTTATTTGCAAGGCGA--GACATTTGTTTTTGGGAAACATTTCCAAATTTCTGT	861
QY	1586	ATATACCCAAATTTGTAGCAGCTATCAAAATGCAAAACATGTTCTCTGGAATGCTAATTTCTA	1645
DB	862	ATATACAGAAATTTGTAGCAGCTATCCAAATGCAAAACATGTTCTCTGGAATGCTAATTTCTA	921
QY	1646	GAGATTGATGACCAATTTTACTTTTCCCAATGCCAGCTATTTAAGAGAAAGGATCAAGG	1705
DB	922	GAGATTGATGACCAATTTTACTTTTCCCAATGCCAGCTATTTAAGAGAAAGGATCAC--GA	979
QY	1706	TGGATTGATGAAGAAGAAAGAAATTAAGCTACAGGGGAGACTAGTTTGCAGTATGTT	1765
DB	980	TGGATTGATGAAGAAGAAAGAAATTAAGCTACAGGGGAGACTAGTTTGCAGTATGTT	1015
QY	1766	ATAATTGATATGAGTCTGTTTGGAAACATTTGATACAACTGGAATTAAGTATGCTTGAAGAG	1825
DB	1016	ATGTTTGTATGAGTCTGACGTGA-----	1041
QY	1826	GTGAAGAAGATTACAGAGAAAGAGCTACAGCTTGTGTTTGGTCAATCTCTGTAAGTAA	1885
DB	1042	-----TGCTTGGGGAGATTCCAAAGAACGAG	1067
QY	1886	GTGATGAAGAACTGAACAAATCGAAGTTCCAAATCATTTAGGGAAGAAATGGATCTAT	1945
DB	1068	GTGATGAAGAACTGAACAAATCGAAGTTCCAAATCATTTAGGGAAGAAATGGATCTAT	1127
QY	1946	CTGACTGTTGAAGAGCGCTTGGAGCATGCAACTTCAATCTACGTGCAAGCAAAACGAAC	2005
DB	1128	CTGACAGTTGAAGAGCGCTTGGTGGTGCATCTTCACTTCCATCTCTTCAAAACCGAAC	1187
QY	2006	CCAAAGAAGATGAACAGAGGTTTGGAAACAAATGTGTGACTGAGTCA	2052
DB	1188	CCCAAGAAGATGAATCAGAGGTTTGGAAACAAATGCTTAATAATAA	1234

RESULT 6

US-10-425-115-50601
; Sequence 50601, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 50601
; LENGTH: 2727
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_146143C.1
US-10-425-115-50601

Query Match 27.7%; Score 677.4; DB 20; Length 2727;
Best Local Similarity 60.8%; Pred. No. 3.6e-171;
Matches 1123; Conservative 0; Mismatches 721; Indels 3; Gaps 1;
QY 122 GTGCCAAGTCCAGGTTTCCACCGCCAGCCGGTTTTTCAAGTCTCTAAAGTACTTCTTTG 181

Db 300 GGGTACAAAGTCGGCGTCCACCGGAGAAACCTCTCTCGCCGAGATCTCTGACGGGTG 359
Qy 182 AAGGAGACTTCTCCCTGATGACCCCTTTGAGCGAGTTCAGAACAGCGCAGCTTCCAG 241
Db 360 AAGGAGACGTCTTCGCGCGAGCCCGCTCGGGAGTACAGGACACCGACCGTCCAAG 419
Qy 242 AAGTTATGCTTGCGCTTTCAGTTCTTCTCCCAATTTTCGAATGGGCTCCCAATACACC 301
Db 420 AAGATCTGGCTCGGCTGAGACATCTTCCCGTGTGCAATGGAGCGCATTAATCC 479
Qy 302 TTTCAGTTCGAAAGCTGACCTCATAGCTGGCATCAACATCGCTAGCTTGGCCATTCCT 361
Db 480 CTCGGCAAGTTCAGGGTGACTTCATCGCTGGGCTACCAATTCGCGCTCTGCATACCC 539
Qy 362 CAGGGCATCAGTTATGCAAGCTCGCAACCTCCCTCCAATTTCTGGACATATTCGAGC 421
Db 540 CAGGACATCGGTATTTAAGCTTGTAACTTTCGACGAGAGGTGGGACTATACAGTAGC 599
Qy 422 TTATATACCACTTGAATTTATGCGATGATGGGTAGCTCGAGGATTTGGCGAGTGGGACT 481
Db 600 TTCGTGGCGCTCTGATATACGAGTATGGGAGCTCCAGGACATAGGCCATCGGTCCA 659
Qy 482 GTGCGGTTGGATCGCTTCTGATGGGTTCGATGTGAGTAATGCCGTTGATCCCAATGAA 541
Db 660 GTGCGGTTGGTGTGCTGTCTCGGTACTCTCTTCAGAAATGAGATCGACCCCAAGACA 719
Qy 542 GACCCAAAGCTTTACCTCACCTGGCTTTCACAGCTACATTAATTTGCTGGTGTGTTTCAG 601
Db 720 CACCCGCTCGAGTACAGGCGCCTAGCTTTCACAGCAACCTTTTTCGCGGGGTCACTCAG 779
Qy 602 GCTGCTTGGGTCGTGTTAGGTTGGGTTGATCGTGATTTCTGTACATGCAACCATTA 661
Db 780 GCAGCCCTCGGATTTTCAGGCTAGGTTTATCATAGATTTCTGTCTAAGTCGCAT 839
Qy 662 ATAGGGTTCAATGGGAGAGCAGCAGCGGTGTGTCTGCAGCAACTTAAATCGATTCTT 721
Db 840 GTCGGATTCATGCGCGCTGCCATCACCAATTCGCGCTTCAGCAGCTGAAAGATTCCTT 899
Qy 722 GGCCTTGAGCAATTCACCCATGAGCTGATATCATATCAGTATGCGCTCTGTTTTCACC 781
Db 900 GGAATTTGCAAACTTCACCAAGAAATCCGATATTTGTATCTGTTTGAATCAGTTTGGGA 959
Qy 782 CAAACT---CATGAGTGGAGTGGAAAGTGTGTGTAGGATGTGTTCTCATTTTCTTC 838
Db 960 AATGTTACCATGGGTGGAATGGCAGCAATCTGATAGGCGACCTTCTCTGGCATTC 1019
Qy 839 CTCCTTAGCAACAAGATACTTCAGCAAAAACGACCAAGGTTTTTTGGGTGTGAGCAATG 898
Db 1020 CTTCTGGTTGCCAAGTACATTGGGAAAAGAAATAAAAGCTCTTCTGGGTGTCGCAATC 1079
Qy 899 GCGCCATGACGTCCGTTATATTGGGAAGTCTCTTGGTTTATTTCACTACGCGCAGAG 958
Db 1080 GCACCTCTCACTTCGGTGATCATATCCATTTTGTGTATCATCACTCTGTCAGATAAG 1139
Qy 959 CAGGTTGGAAGTATAGAGAACTGAAGAGGTTTGAATCCACATCACTCACAAT 1018
Db 1140 CATGGCGTTGCAATTTGCAAGACATAAGAAAGGATCAACCCACTTCAGCTAGTCTA 1199
Qy 1019 CTGGTATTTGTGTCGCTTACATGACTACAGCTGTCAAACTGGCATTTGCTGGCATC 1078
Db 1200 ATATACTTACCGGCCATATCTTGGCAGCAGGATTCAAAATTTGGATAGTAGTGAATG 1259
Qy 1079 ATATCATCTTGGGAAGAAATAGCAGTAGGAAGCTTTGCAATGTATATAAATTAATCAAT 1138
Db 1260 ATAGGCTTAACGGAAGGATTTGCAATTTGGAAGAACATTTGCAGCCCTCAAGGATACCG 1319
Qy 1139 ATTGATGGCAAAAGAGATGATAGCTATTTGGACCATGACAGCTAGTTGTTTTCACC 1198
Db 1320 ATAGATGGGAACAAAGAAATGATGGGCTCTAGGACCATGAACATTTGTTTCACTGACT 1379
Qy 1199 TCTTGCTACTTCACAAAGGACCATTTTCGGTTCGGCTGTGAACTATTAACGCTGGATGC 1258

Db 1380 TCTTGCTAGTAGCCACAGGTTCTTTTCTCGCGTCAGCAGTTAAATTACATGGCTGGCTGC 1439
Qy 1259 AAGACAGCAGCTTCCAAATTAATATGTCACTTGTCAGTAAATGTTGACATTTGTTATCCCTG 1318
Db 1440 AMAACAGCAGTGTCAACCGTGTGTTATGTAATTTGTCGTAAATGCTTACTTTGCTACTGATC 1499
Qy 1319 ACACCCCTTGTTCATTTACACTCCCTCGTGGTGCTATCATCAGCTATTAATCGTATCTGCAATG 1378
Db 1500 ACCCATTTGTTCAAGTACACTCCAAATGCCATCTTCTCGATCATCATATCAGCAGTG 1559
Qy 1379 CTTGGACTCATAGATTAAGAGCAGCCATCATCTATTAAGGTTGACAAATTTGACTTT 1438
Db 1560 CTTGGTTAAATGACTATGAATCGGCTTACTTATCTGAAAGTTGACAACTGGACTTT 1619
Qy 1439 GTGGTGTGATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTTCGAAATTTGCTTAGTC 1498
Db 1620 CTAGCATGATGGAGCATTTTTTGGAGTCATATTTTCATCGGTGAGTATGGCTTGCTC 1679
Qy 1499 ATAGCTATTTAATATCTGTACTTCGGGTACTTCTATTTATTTGCAAGGCCAAGCAGATTC 1558
Db 1680 ATTGCGTTGCAATATCTCTTCTCAAAATTTCTTCCAAGTAACACGCGCAAGACAGTT 1739
Qy 1559 GTTTGGGCAACATTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCA 1618
Db 1740 TTACTTTGAAACCTTCCAGAACCACTATATACAGGAATGTAGAACAGTATCTGATGCT 1799
Qy 1619 AAACATGTTCTCGAATGCTAATTTCTAGAGATGTAGACCAATTTACTTTGCAATGCC 1678
Db 1800 ACCAAGTTCCAGGGGTGCTGATTTGTAGAGTGATCTCAGCTATATACTTTCACAACTCT 1859
Qy 1679 AGCTATTTAAGAAAGGATCAAGGTGATGATGAAGAAAGAAAGAAATTAAGCT 1738
Db 1860 AACTATGTTAAGAGAGATCTGAGTGTCTAAGAGATGAGGAGGACCAACAGGAC 1919
Qy 1739 ACAGGGAGACTAGTTTGCAGTATGTTATAATTTGATATCAGTGTCTTTGGAAACATGAT 1798
Db 1920 CAGAAGTTAACAAAAACCTGAGTTTCTAAATTTGTGACCTGTCTCTCTGTAATGATATCGAC 1979
Qy 1799 ACAGTGGAAATGATGCTTGAAGGTTGAAGAGATTTACAGAGAGAGAGAGCTACAG 1858
Db 1980 ACAAGTGGAAATCAATGCTTTGGAGGAGTTGGGAAAGCTCTTGAAAAACGCAAAATTCAG 2039
Qy 1859 CTTGTTTTGCTCAATCTCTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAA 1918
Db 2040 CTGTTCTTACCAATCCCGGCGCGGTGATCCAGAGCTCCGCTCAGGAAATTCAG 2099
Qy 1919 AATCATTTAGGAAGAAATGGATCTATCTGACTGTTTGAAGAGCGCT 1965
Db 2100 GACATGTTGGTGAAGACAAACATATTCCTCAGCGTCGGCGAGCTGT 2146

RESULT 7

US-10-437-963-77012
; Sequence 77012, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Roba, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77012
; LENGTH: 4390
; TYPE: DNA

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 65103
; LENGTH: 2366
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29800C.1
US-10-424-599-65103

Query Match 26.9%; Score 658.6; DB 18; Length 2366;
Best Local Similarity 60.5%; Pred. No. 3.9e-166;
Matches 1120; Conservative 0; Mismatches 724; Indels 7; Gaps 2;

QY 77 AGTGTAGATATGAGTACCTTTGGGCATGAACAACTTTGAGAGATGCGACCAAGTCGAG 136
DB 132 AATGCATGTAATGATCAATCTCATATGATAGATGTCATGGAAGTCACCAAGTTGTG 191
QY 137 GTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACTTTCTTC 196
DB 192 CCACGACCTCACAGAGACCCTTCAGAACTCAAGGGTAGGCTCAAGGAACTTTCTTC 251
QY 197 CTTGATGACCTTTGAGGCGAGTTCAAGAAACAAGCCAGCTTCCAAAGAGTTTCATGCTTGC 256
DB 252 CCTGATGATCCTCTGCGCCAAATTCAGGGACACACCTCTTAAAGAGAAAACTGATCCTTGA 311
QY 257 CTTGAGTTCTTCCCAATTTTGAATGGGCTCCCAATATACACCTTTGAGTTCTTGA 316
DB 312 GCTCAATATGTTCCCTATTTCTCAATGGGGTCTTAAGTATAATCTCAACTCTTTCAA 371
QY 317 GCTGACCTCATAGCTGGCATCACATCGCTAGCTTGGCCATTCCTCAGGSCATCAGTTAT 376
DB 372 TCTGACCTGTTTCTGGCCCTCATAATGCTAGCTTGGACATCCCGCAGGGAATGAGTTAT 431
QY 377 GCCAAGCTGCCAACCTCCCTCAATTTTGGACTATATTTGAGCTTTTATACACCAATG 436
DB 432 GCTTAGCTTGCAGTCTTCTCCAAATTTAGGACTTTATCTAGTTTGTTCACACCTT 491
QY 437 ATTTATGCGATGAGGTAGCTCGAGGATTTGGCAGTGGGAGCTGTGGCGGTGATCG 496
DB 492 GTCTATGCTGTTCTTGGAGCTCAAGAGCTTTGCAGTTGGACCTGTTCTATTTGCTTCT 551
QY 497 CTTCTGATGGTTCGATGTTGAGTATGCGTTGATCCCAATGAAGACCCAAAGCTTTAC 556
DB 552 CTTGTATGGATGCCATGTTGCATCAGGAAGTGTCTCCCAACACAGATCCAAATCTGTTT 611
QY 557 CTCACCTGGCTTTACAGCTACATTAATTTGCTGGTGTGTTTTTCAGGCTGCTTGGGCTG 616
DB 612 CTTGAGCTAGCTTTCACTTCAACATTAATTTGCGGTCTCTTTCAAGCTTTGCTTGAATC 671
QY 617 TTTAGTGGGTTGATCGTGGATTTCTGTCACTGCAACCAATATAGGGTTCAATGGA 676
DB 672 CTAAGGCTAGGCTTCATAATTTGATTTCTATCTAAGGCCATCTTATTTGGGTTTCAAGCT 731
QY 677 GGAGCAGCCACGGTGTGTCTGAGCAACTAAATCGATTTCTTGGCCCTTGAAGCTTTC 736
DB 732 GGAGCTGCTATTAATTTCTCACTGCAACAGCTCAAGAGCTGCTTGGAAATCAACATTTTC 791
QY 737 ACCCATGGAGCTCATATCATATCAGTGATGCGCTCTGTTTTTACCCAAACTCATGAGTGG 796
DB 792 ACTAATCAGATGGGCTGATCTCTGTATGACTTCTGTTTTTCAACAATATACATGAGTGG 851
QY 797 AGGTGGAAAGTGTCTGTTAGGATGTTCTTCATTTTCTTCTCTCTTGAACAAGATAC 856
DB 852 TCATGGCAACAATAATTGATGGGGAATTTGCTTCTTGGTACTACTACTATTATAGCAAGAC 911
QY 857 TTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTAGCAATGGCGCCATTTGAGCTCGTT 916
DB 912 GTTAGCATAGGAAACCAAAACTTAATCTGGGTCTCAGCTGGAGCTCCTCTTATGTGTGTC 971
QY 917 ATATTGGGAAGTCTCTTGGTTTATTTCACTTCAGCCGAGAGCAGCGTGTGAAGTGATA 976

RESULT 9

US-10-425-115-182207
; Sequence 182207, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

DB 972 ATCATCTCTACCTCTCTGGTTTTTGCAAATTAAGGCTCAAAATCATGGCATCAGTCGATTT 1031
QY 977 GGAGAACTGAAGAAGGGTTTTGAATCCACCATCACTACAAAATCTGTATTTGTGTGCGCT 1036
DB 1032 GGAAAATTCACAAGAGAAATAATCTCCATCATGGAATATGTTGCTCTTTTCATGGAAGT 1091
QY 1037 TACATGACTACAGCTGTCAAAACTGGCATTTGCTTTGGCATCATATCACTTGC - GGAAGG 1095
DB 1092 CACCTAGGCTAGTTTATGAAAACAGGGCTTATACCGGATTTTGTCCCTAAACAGGAAG 1151
QY 1096 AATAGCAGTAGGAAGAAGCTTTGCAATGATATAAAAATTAACAATTTGATGGCAACAAAGA 1155
DB 1152 TATTCAGTAGGAAGACATTTTGCAGCTCTCAAAAACCTCAAAAAGGATGGAATTAAGGA 1211
QY 1156 GATGATAGCTATTTGGGACCATGAACGTAGTTGGTTCTTTCACTCTTGTCTACTCAAC 1215
DB 1212 AATGATGGCAATTTGGTTTTATGAAATTTGTTGGCTCTCTTCACTTCTGCTATGTTACAAC 1271
QY 1216 AGGACCAATTTTCGGCTTGGCTGTGAACCTATAACCTGGATGCAAGACAGAGCTTCCAA 1275
DB 1272 AGTGCTTCTCTCGTCAAGCTTTAAACAACATGACGCGCAAAAACAGCTGTGTCAA 1331
QY 1276 CATTAATATGCTCACTTGCAGTAATGTTGACATTTGTTATTCCTGACACCTTGTTCATTA 1335
DB 1332 TGTAGTGTGCTGTGACAGTCATGTCACCTCTCTTTTCTCATGCCATTTGTTTCAATA 1391
QY 1336 CACTCCCTGGGTGCTATCAGCTATATCGTATCTGCAATGCTTGGACTCATAGATTA 1395
DB 1392 CAGGCTAATGTCGTGGGCGCAATCATAGTCACAGCAGTAATTTGGCTCATGCTAT 1451
QY 1396 TGAAGCAGCATCCATCTATTTAAGGTTGACAAAATTTGACTTTTGGTGTGATGAGTGC 1455
DB 1452 CCCCCTGCTGTAACTTTGGAAGATCGACAAATTCGATTTTGTGTGATGACTGC 1511
QY 1456 ATACATTTGGCGGTCTTTTGGCAGTGTGAAATTTGGCTTTAGTCACTAGTATTTGTAATC 1515
DB 1512 TTTCTTAGGTGTTCTTTTCTCTCTGTCAGGAGGCTTGTCTTGTGTTGGATATC 1571
QY 1516 TGTACTTGGGTACTCTTATTTATTCAGGCAAGGACATTCGTTTGGGCAACATTC 1575
DB 1572 AACTTTGAAGATACCTTTGCAAAATTCAGAGACCGAAACAGTGATTTGGGGAAGATACC 1631
QY 1576 AAATTTCTGTATATACCGAAATGTTGAGCATCATCAAAATGCAAAACATGTTCTCTGGAAT 1635
DB 1632 AGGAACACATATATAGAAATCTTGATCAATACAGGAGCTGTGAGATACCTGGAT 1691
QY 1636 GCTAATCTTAGAGATTGATGCAACCAATTTACTTTCCCAATGCCAGCTATTTAAGAGAAAG 1695
DB 1692 TCTTATTTTAAGCAATTGAGGCTCCCATCAATTTTGCACATCATATCTCAACGAGAG 1751
QY 1696 GATCAAGGTGGATTGATGAAGAAGAAGAATTAAGCTACAGGGGAGACTAGTTT 1755
DB 1752 AACGTTAAGATGGATTGAAGAAGAAGAACAATAA-----GGAACAAATTAAGCCT 1805
QY 1756 CGAGTATGTTATAATTTGATATGAGTGTGTTGGAACAATGATACAAGTGAATTAAGTAT 1815
DB 1806 TCGATTTCTTAGTTTGAATTTGCAATTTGCTGAGTGTGAGTGTGTTGACACAGTGGATCTCACT 1865
QY 1816 GCTTGAAGAGGTGAAGAAGATTTACAGAGAGAAGAGAGCTACAGCTTGTGTTGGTCAATCC 1875
DB 1866 TTTCAAGGAGTTGAAAGCAACACATCGAAAAGAAAGGTGTTGAGCTTGTGTTGGTCAATCC 1925
QY 1876 TGTAAAGTGAAGTGAAGAAACTGAAACAAATTCGAAGTTTCCAAAATCATTT 1926
DB 1926 TCTTCTGAGGTTCATAGAAAAGCTTTAAAAAAGCAGATGAAGCTAATGATTT 1976


```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 182207
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRF4577_97750C.1
US-10-425-115-182207

Query Match          26.6%; Score 650.6; DB 20; Length 2651;
Best Local Similarity 60.2%; Pred. No. 6e-164;
Matches 1117; Conservative 0; Mismatches 729; Indels 9; Gaps 2;

QY 121 AGTGCAACCAAGTCGAGGTTCACCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTT 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GAAGGAGACTTCTTCCCTGATGACCCCTTGAGGCGAGTTCAGAAACAAAGCCAGCTTCCAA 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 577 GAAGGAGAGACTTCTTCCCGGACGACCCGTTCCGGGGGTTTCAAGGGGCGAGCCCGGGGAC 636
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 GAAGTTTCATGCTGGGCTTTCAGTTCTTCTTCCCATTTTCGAATGGGCTCCCAATACAC 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 637 GCAGTGGCTCATGGCGGTGAGGTACCTCTTCCCATCTTGGACTGGGTCCGAGCTACTTC 696
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 CTTTTCAGTTCTTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCC 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 697 CTTGTGCTCTTCAAGTCCGACCTCGTTCGCGGSCCTCACCATTCGCCAGCTCGCCATTCC 756
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 TCAGGCGCATCAGTTATGCGAAGCTCGCCAACTCCCTCCCAATTCCTTGGACTATATTCGAG 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 757 TCAGGCGCATTTAGCTACGCGAAGCTGGCAAGCTTGGCTCCCAATTCGGGCTGTATTTCGAG 816
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 CTTTATACCAACCATTCATTTATGCGATGATGGTAGCTCGAGGATTTGGCAGTGGGAC 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 817 CTTTGTGCGCCGATGGTGATCGCGGTGCTGGGGAGCTCCCGTGACCTGGCGGTGGGCC 876
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 TGTGGCGGTTGGATCGCTTCTGATGGGTTCCGATGTTGAGTAATGGCGTTGATCCCAATGA 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 877 GGTGTGATCTCGTCTGCTGATCATGGGTCATGCTGCGGCGAGCGGTGAGCCCGC 936
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 AGACCCAAAGCTTTACTCCACTCGGCTTTTCAGACTACATTTATTTGCTGGTGTTTTCA 600
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 937 GGAGCCGAGCGCTTCTCTGAGCTGGGCTTTCACCTCCAGCCCTGTTCCGCGGGGTGGTGA 996
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 GGTGCTCTTGGGCTCTGTTAGGTTGGGTTGATCGTGGATTTCTGTCACATGCAACCAT 660
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 997 GGCTCTCTTGGGATCCTCAGGCTCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 AATAGGGTTTCATGGGAGGAGCAGCCAGGTTGGTGTGCTGCGACCACTAAATTCGATTCT 720
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1057 GGTGGGTTTCATGGCGGCGCCCATCATCTGTTGGGCTGCGAGAGCTCAAGGGCTGCT 1116
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 TGGCTCTGAGCAATTTCAACCATGGAGCTGATATCATATCAGTGAATGGCTCTGTTTTCAC 780
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1117 GGGCATCGTCCACTTCACCAACGAGATGGGATCGTCCAGTTCATGGGCTCGGCTCTTCCA 1176
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 781 CCAAACTCATGAGTGGAGGTGGAAAGTGTGTTAGGATGCTCTTCAATTTCTTCTCT 840
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1177 CCACACAGCGAGTGTGCTGGCAGACGATCTCTCATGGGCGTCTGCTCTCTGCTCTCTCT 1236
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 841 CCTTAGCACAAGATATCTTCAGCAAAAAAGACCAAGGTTTTTTTTGGGTTTCAGCAATGGC 900
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1237 GCTGTGCGGCGAGCATGTGAGCATCAGATGGCCAAAGCTTTCTGGGTTTCGGCGTGGC 1296
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 901 GCCATTGACGTCGTTATATTGGGAAGTCTCTTGGTTTATTTCATCTACGCGCGAGAAGCA 960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1297 GCGCTGTCATCGGTGTCACCATCTCGACGCTGCTTGGTTTCTCTTCAAGCTCAGAACA 1356
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 CGGTGTTGAAGTGATPAGGAGAACTGAAGAAGGTTTGAATCCACATCACTACAAATCT 1020
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1357 TGGCATCAGCATCATTTGGGCGAGCTCAAGTGGGCGCTGAATCGCCCTCTGTTGGACAAGCT 1416
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 GGTATTTGTGTCGCTTACATGACTACAGCTGTCAAAACTGGCATTTGTTGGCATCAT 1080
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1417 CCTGTTTGACACGGCGTATTTAGGCGCTCACCATGAAGACTGGCTTGTCAACGGAAATCAT 1476
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1081 ATCACTTCGCGAAGGAATAGCAGTAGGAAGAGCTTTTCAATGTATAAAAAATTAACAATAT 1140
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1477 CTCACCTGACGAAGGAATAGCGTTGGTGAACAATTTGCCCTCACTCAGGAGCTACAGAT 1536
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1141 TGATGGCAACAAAGAGATGATAGCTATTTGGGACCAATGAACGTAAGTGTGTTTTCACCTC 1200
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1537 AGATGGAAACAAGGAGATGATGGCCATAGGTTGTATGAATGTTGTGGGTCTCGACATC 1596
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1201 TTGCTACCTCACAACAGGACCAATTTTCGGTTCGGCTGTGAACCTATAACGCTGGATGCA 1260
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1597 ATGCTACGTAACAACAGGTGCGTTCTCCGCTCTGCTGTAAACCAACACGCGCGCTGCAA 1656
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1261 GACAGCAGCTTCCAAACATTTATAATGTCACTTTCAGTAATGTTGACATTTATTCTCTGAC 1320
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1657 GACTGCCATGTCACAGTGATCAAGGCGCTGACTGTGATGTCAGCTGCTGTTCTCTCAT 1716
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1321 ACCCTTGTTCATTACACTCCCTCGTGGTGTGCTACAGCTATTAATGCTATCTGCAATGCT 1380
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1717 GCCACTGTTCTGTATACACACCAACGTTGCTCTCGAGCGATCATCATCGCGCGGTGAT 1776
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1381 TGGACTCATAGATTAAGAGAGCCATCCATCTATTAAAGTTGAACAATTTGACTTTGT 1440
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1777 CGGCTGATCGATTTTCCCGCGGTGTACACATCTGGAAGATGGAACAAGATGGAATTTTCT 1836
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1441 GGTGTGATGAGTGATACATTTGGGCTGTCTTTGGCAGTGTGAAATTTGGCTTGTAGTAT 1500
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1837 GGTGTGCGTTTGGCGGTTTGGCGGCTCATCTTATCTCAGTCCAAAGAGCGCTTGGCAT 1896
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1501 AGCTATTCTTAATATCTGTAATTTCCGGTACTTCTATTTATTCGACGGCCAAAGACATTCGT 1560
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1897 AGCGGTGTGATATCTATATTTAGGCTGTGTATGTCAGATCAGAGGCCCAAGATGATGT 1956
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1561 TTTGGGCAACATTCCAAATTCGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAA 1620
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1957 TCAAGGGAACATCAAGGGGACTGATATTTTACAGAGACCTGTCATCTCAAGGAGGCCCA 2016
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1621 ACATGTTCTCGAATGCTAATTTCTAGAGATTGATGCACCAATTTTACTTTGCCAATGCCAG 1680
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2017 AAGAGTTTCTGGGTTCTTGATCTTGGCCATTGAAGCACCGATAAACTTTCGCCAATCCCA 2076
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1681 CTATTTAAGAGAGAAAGGATCAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTAC 1740
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2077 CTACCTGATGAAGGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2130
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1741 AGGGGAGACTAGTTTGCAGTATGTTATTAATGATATGAGTGTGTTGAAACATTTGATAC 1800
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2131 TAAACATACTGAACTCCATTTTCATAATCTTGGATCTGTGAGTGTCTCTGCAATTTGACAC 2190
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1801 AAGTGGAAATAGTATGCTTTGAAGGTTGAAGGATTAACAGAGAAAGAGAGTACAGCT 1860
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2191 AAGTGGCATAGGTTTCTTCAATTTGACATTAAGAAATCAATAGAGAAACGTTGTTGGAGCT 2250
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1861 TGTGTTTGTCTCAATCTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1917
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2251 TGTGCTTGTCAATCAACTGGGAGATCATGGAGAAATACAACTGTCGAAACGAGGCTGA 2310
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1918 AAATCATTTAGGGAAGAAATGATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1972
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2311 AAATCATTTTAGGCCAGATTTGCTTGTATCTGACCACTGGCGAAGCAATGCTTCA 2365
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


RESULT 10

US-10-437-963-77005

Sequence 77005, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 77005

LENGTH: 2780

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_76946C.1

US-10-437-963-77005

Query Match 26.4%; Score 647.2; DB 19; Length 2780;

Best Local Similarity 59.8%; Pred. No. 5.1e-163;

Matches 1103; Conservative 0; Mismatches 738; Indels 3; Gaps 1;

Qy 125 CACCAAGTCGAGGTTCCACCGCCACACGCGTGTTCCTCAAGTCTCTAAAGTACTCTTTGAAG 184

Db 457 CACAAGGTCGAGTTCCTCCACCGAAGAAGACTCATCGACGAGTTACCGACGCGGTGAAG 516

Qy 185 GAGACTTTCTCCCTGATGAGACCCCTTTGAGGCGAGTTCAAGAACAAAGCCAGCTTCCAAAGAAG 244

Db 517 GAGACGTTCTTCGCGCGACGACCCCGTTGCGGCAGTACAAAGACGACGCGATGTCAAAGAAG 576

Qy 245 TTCATGCTTGGCTTCAGTCTTCTTCCCATTTTCGAATGGCTCCCAATACACCTTT 304

Db 577 GTGTTGATCAGCTTCGCAAACTTCTTCGCGTGTGAGTTCGGGCGCGGACATACACCTTC 636

Qy 305 CAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAG 364

Db 637 CGAAGTTCAGGGCGACCTCGTCTCGGCTCACCATTCGCCAGCTCTGCTATACCCAG 696

Qy 365 GGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATTCCTTGGACTATTCGAGCTTT 424

Db 697 GACATCGTTATGCGAAGCTTGTGCGCTATTACCAAAATATATGGAAGTGTACAGCAGCTTC 756

Qy 425 ATACCACCATGATTTATGCGATGATGGTGTAGTGTGAGGATTTGGCACTGGGACTGTG 484

Db 757 GTACCGCTTTGATATACGCGATGATGGAAGTTCCAGGGATATAGCCATTTGGTCCAGTG 816

Qy 485 GCGGTTGATCGCTTCTGATGGGTTTCGATGTTGAGTAATGCGGTTGATCCCAATGAAGAC 544

Db 817 GCGTGTGTTGCTGCTGCTTGGCACTCTCCTCCAGATGAGTTGATCCCAAGAGAT 876

Qy 545 CCAAGCTTTTACCTCCACTGGCTTTTCAAGCTACATTAATTTGCTGGTGTTCGAGCT 604

Db 877 CAGGAAGAGTACACCCGCTGGCTTCACTGCAACGTTCTTCGCTGGGCTCACCCAGGCA 936

Qy 605 GCCTTGGTCTGTTAGGTTGGGTTGATCGTGGATTTCTGTCAATGCAACCAATAATA 664

Db 937 GTGCTCGGATTTCTCAGGCTAGGGTTTATATAGAGTTCTTGTCCCATCTGCCATCGTT 996

Qy 665 GGGTTTCATGGGAGGAGCAGCAGCGTGGTGTCTGTGACCAACTAAATCGATTTCTTGC 724

Db 997 GGATTCATGGGAGCGGCGCATACCAATGGCTTCAGCAGCTTAAAGGCTTCCTTGA 1056

Qy 725 CTTGAGCATTTTCAACCATGGAGCTGATATCATATCAGTGTGCTGTCTTTTACCCAA 784

```
QY 1862 GTTTTGGTCNATCCTGTAAGTGAAGTGAAGAACTCAACAATCGAAGTCCAAAT 1921
Db 2197 ATTCTAGCCAAATCCGGGGCCAGCTGTGATCTCTGAAGCTCCGATACGGAAATTCACGGAT 2256
QY 1922 CATTTAGGGAAGAAATGGATCTATCTGACTGTGTGAAGAGGCGGT 1965
Db 2257 CTCATCGGTGAAGACAAGATATTCCTGACTGTCTGGGGACGCCAGT 2300

RESULT 11
US-10-437-963-95853
; Sequence 95853, Application US/10437963
; Publication No. US200401233431
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 95853
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94005C.1
US-10-437-963-95853

Query Match 24.5%; Score 600.6; DB 19; Length 2001;
Best Local Similarity 58.2%; Pred. No. 1.5e-150;
Matches 1094; Conservative 0; Mismatches 779; Indels 6; Gaps 2;

QY 145 GCCACAGCGGTTTTCAGAGTCTCTAAAGTACTCTTTGAAGGAGAGCTTTCTCCCTGATGA 204
Db 123 GCGCGCGCGTTTCGGGGAAGCTGTGTGCGAGCTTCGCGGAGAGCTTTCTCCCGCGAGA 182
QY 205 CCCTTTGAGCGAGTTCAAGAACCAAGCCAGCTTCCAAAGAGTTTCATGCTTGGCTTTCAGTT 264
Db 183 CCGGTTCCGCGGTTTCGGCGGCTGCCCGCGCCAGGCGCGGTGTGCGCGGTCAAATA 242
QY 265 CTTCTTCCCAATTTTCGAATGGGCTCCCAATACACCTTTGAGTTCTTTGAAGCTGACCT 324
Db 243 CTTCTGCTCCGCGGCTGGAAGTGGGTCGCGGCTACGCGCTTCGACAAAGTTCAAGTTTCGACCT 302
QY 325 CATAGTGCATCACCATCGTAGCTTGGCCATTCCTCAGGCGCATCAGTTATGCCAAGCT 384
Db 303 CTTCTGCGGCGATCACCATCGCAGCTTCGCCATTCGCGAGGCGATCAGCTACGCCGCGCT 362
QY 385 CGCCCAACCTCCCTCCAAATTTCTTGGACTATATTTGAGCTTTTATACCAACCATTTGATATGC 444
Db 363 CGCCCAACCTCCCTCCAAATCATCGGCTCTATTCGAGCTTCGTGCGCGCGGTGATGACG 422
QY 445 GATGATGGTAGTCTGAGGGATTTGGCAGTGGGAGCTGTGGCGGTTGGATCGCTTCGAT 504
Db 423 GGTGTTTCGGGAGCTCCAAACACTGCGGTTGGGAGCGGTGGCGCGGCTCGCTCGCT 482
QY 505 GGGTTCGATGTGAGTAAATGCGGTGATCCCAATGAAGACCCAAAGCTTTTACCTCCACT 564
Db 483 GCGGTCATCATCGAGAAGAGTGGCGGAGCGGAGCAACCGAGCTGTACTCGAGCT 542
QY 565 GGCCTTTCACAGCTACATTAATTTGCTGGTGTGTTTTTTCAGGCTGCTGGGTTCTGTTAGGTT 624
Db 543 CTTCTACACCGCGCTTCTTCACCGGCTCTTTCAGAGCGCGCTAGGCGTCTTTAGGTT 602
```

```
QY 625 GGGGTTGATCGTGATATTTTCTGTCACTGCAACCATTAATAGGGTTCATGGAGAGCAGC 684
Db 603 AGGCTAATAGTGAATTTCTCTCGCGTCAAGCATCACCGGTTTCATGGCGGCACGCG 662
QY 685 CACGGTGGTGTCTGAGCAAACTAAATCGATTTCTTGGCCCTTGAGCATTTTCAACCATGG 744
Db 663 GATGATCATCATCTGAGCAGGTTCAAGGGGCTCTCGGGATGAAGCACTTTCACCAACCA 722
QY 745 AGCTGATATCATATCAGTGAAGCGCTGTGTTTCAACCCAACTCATGAGTGAGGTGGGA 804
Db 723 GACAGACATCATCTCCGTCCTGCACTCACCTACATTAACCGGACAGGTGGAATGSCA 782
QY 805 AAGTGTGTGTAGGATGTGCTTCAATTTCTTCTCTTAGCAACAAGATATTTCAAGCAA 864
Db 783 GAGTGCAGTTCTCGGCATATGCTTCTCTTGTCTTGATGTCGAGTAAGCACCTGAGAA 842
QY 865 AAAACGACCAAGGTTTTTTTGGGTGTGAGCAATCGCGCATTTGACGTCGCTTATTTGGG 924
Db 843 GAAACTGCCAAAGTTGTTCTGGGTGTGCGCATTTGCGCCATTCATGTTGTCGTCATCGG 902
QY 925 AAGTCTCTTGGTTTTATTTCACTCAGCGCGAGAGCAAGGTGTTGAAGTATAGGAAGCT 984
Db 903 AGGCATCTTCGCTTTCTTGTCAAGGGTGAAGCATGGAATCCGATAGTCGGTAACCT 962
QY 985 GAAGAAGGGTTTGAATCCACCATCACTCAAAAATCTGGTATTTGTGTGCGCTTACATGAC 1044
Db 963 GAATTAAGGATCAATCCCTGTCCATTTGCGCAGCTAACGTTCCAGTCCAGGCACATGAA 1022
QY 1045 TACAGCTGTCAAACTGGCAATTTGCTTGGCATCATATCACTTCCGGAAGAAATAGCAGT 1104
Db 1023 GACGCGAGTGAAGCTGTGCTCTTGTCTGGGATCCTAGCACTGGCAGAGAATAGCCGT 1082
QY 1105 AGGAAGAGCTTTGCAATGTATAAATTAATTAATTTGATGGCAACAAGAGATGATAGC 1164
Db 1083 CGGCGGAAGCTTGGCGATGTCGAAGACGAGCAGATCGACGCGCAACAGGAGATGATCGC 1142
QY 1165 TATTGGGACCATGAACGTAGTTGGTCTTTTCACTCTTGTCTTCACTCAACAAGACCACTT 1224
Db 1143 GTTTGGTATCATGAACATCGCGGTTCTTTACCTCTGCTACCTCAGCAGAGGCGGTT 1202
QY 1225 TTCGCGTTGGCTGTGAATATAACGCTGGATCAAGACAGCAGCTTCCAAATATTAAT 1284
Db 1203 CTCAAAGTCGCGGCTGAACTTCCACGCGGGTCAAGACGCGGATGTCCAAACGTGGTGAT 1262
QY 1285 GTCACTTGCAGTAATGTGACATTTGTTATTTCTGACACCTTGTTCATTAACATCCCT 1344
Db 1263 GTCGGTGTGATCATGCTGTGCTGCTGCTTCTGGCACCTCTTCAAGTACACCCGCT 1322
QY 1345 GGTGTGCTATCAGCTATTATCGTATCTGCAATGCTTGGACTCATAGATTATGAAGCAGC 1404
Db 1323 GGTGGCGCTCTCTCCATCATCTGTCGCCATGATCGGGCTCGTCAAGGTCNAGGAGTT 1382
QY 1405 CATCCATCTATTTAAGGTTGACAAATTTGACTTTGTGTGTGTCATGAGTGCATACATGG 1464
Db 1383 CTGCGCACCTCTACAAGGTCGACAAAGTTCGACTTCTGCACTCTGCACTGTGTCGCTTCTCGG 1442
QY 1465 CGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCACTAGCTATTGTAATCTGTGACTTCG 1524
Db 1443 CGTGTCTTCTTCCATGCTGCGAGGCTCAGCGCATCCGTAAGCTTGTGAGTTGTGAG 1502
QY 1525 GGTACTTCTATTTATTGCAAGGCCAAGGACATTTGTTTTGGGCAACATTCCAAAATTCGT 1584
Db 1503 GCGGTTGTGTGAGTGGCCAGGCTGCCACCTGCAAGCTCGGGAACATAGCAGGAGCGA 1562
QY 1585 GATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATGTTCTGGAATGCTAAATTC 1644
Db 1563 GACCTTCGCGACGTGAAGCAATACCCCAAGCGAAGAGCATCCCTGSCATCTCTGCTT 1622
QY 1645 AGAGATTGATGCAACCAATTTACTTTGCAATGCCAGCTATTTAAGAGAAAGATCACAAG 1704
Db 1623 GCAGCTCGGCTCTCCCATCTACTTGTGTCGTTACCTGCGGAGAAAGATTTGAG 1682
QY 1705 GTGGATTGATGAAGAAAGAAAGAAATTTAAAGCTACAGGGGAGACTAGTTTTCAGTATGT 1764
```

```
Db 1683 ATGG--GTGGAAGATGAGGATAACCTGTGCAAGAGCGTCGGGACGATCTGCAATACTT 1739
QY 1765 TATAATTGATGATGAGTGTCTTTGGAACATTTGATACAAAGTGGGAATAAGTATGCTTGAAGA 1824
Db 1740 GGTCTTGATCTCTGGTGTGTCTCTGTCGCAACTCTGGAGTTGGGATGCTACTAGA 1799
QY 1825 GGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTTGTGTTTGGTCAATCCTGTAAAGTGA 1884
Db 1800 AGTACACAGAGCTCGAACGAAGAGGATCACGATAGCTCTGACGAATCCGAGGCTAGA 1859
QY 1885 AGTGTGAGAAACTGAACAATC---GAAGTCCAAATCATTTAGGGAGAAATGGAT 1941
Db 1860 GGTAAACAGAGAGCTGTGTCTGTGATACGTACGGGACATCTTAGGGGATGAGTGGGT 1919
QY 1942 CTATCTGACTGTTGAAGAGCGGCTTGAGGATCAACTTCAATCTACGTGCAAGCAAAAC 2001
Db 1920 CTTCTGACGCTCAAGAGCGCCATCAGGCGTGTGATAGCGGCTGCGATATCCAGAAA 1979
QY 2002 GAACCCAAAGAAAGATGAA 2020
Db 1980 TRAGGGAGAGACGAAGTA 1998
```

RESULT 12

```
US-10-437-963-73410/c
; Sequence 73410, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ IDS: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 73410
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73694C.1
US-10-437-963-73410
```

```
Query Match 23.3%; Score 571.4; DB 19; Length 2429;
Best Local Similarity 58.1%; Pred. No. 1.3e-142;
Matches 1075; Conservative 0; Mismatches 726; Indels 48; Gaps 2;

QY 123 TGCACCAAGTCGAGGTTCCACCCGACAGCGCTTTTCAAGTCTCTAAAGTACTTTTGA 182
Db 2260 TGCAAAAGTGTCTGTTGCCGAGAGGAGGTCGACGGCGAAGGCGCTGAGCGCGCTGG 2201
QY 183 AGGAGACTTTCTTCCCTGATGACCCCTTTGAGGCGAGTTCAAGAACAAAGCCAGCTTCCAAGA 242
Db 2200 CGGAGGTGTTCTTCCCGGACGCCGCTCCACAGTTCAAGAACAGTCTGCGCGCGC 2141
QY 243 AGTTCATGCTTGCCCTTCAGTCTTCTTCCGATTTTGCATGGGGCTCCCAATACACCT 302
Db 2140 GGTGTGTGCTCGCGTGCAGTACTTCTTCCCATCTTCCCATCTTCCATTTGGGCTCCGACTACAGCC 2081
QY 303 TTCAGTCTTGTGAAGCTGACCTCATAGCTGGCATCACCATCGTAGCTTGGCGATTCTCTC 362
Db 2080 TCCGCTCTCTCCGCTCCGAC----- 2061
QY 363 AGGGCATCAGTTATGCAAGCTCGCAACCTCCCTCCAAATTTCTTGACTATATTCGAGCT 422
```

```
Db 2060 --GGAATCAGCTACGCCAAGCTCGCAACCTCCCTCCAATCATTGGACTATATTCGAGCT 2003
QY 423 TTATACCAACCATGATTTATGCGATGATGGTAGTCTCGAGGATTTGGCAGTGGGACTG 482
Db 2002 TCGTGGCGCGCTGATCTACTCTGTTGGTAGTCTCGCGGATCTGGCTGTAGGGCCGG 1943
QY 483 TCGCGGTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTTGATCCCAATGAAG 542
Db 1942 TGTGATAGCTGCTGCTGGTGTGATGGGTTCGATGCTCCGCGAGCGGTGTCGCGGACCAG 1883
QY 543 ACCCAAAGTTTACCTCCACCTGGCTTTACAGCTACATTTATGCTGCTGTTTTCAGS 602
Db 1882 AGCCCATCTCTACCTCCAGCTCGCTTCACTCCACCTCTCTCGCGGCGTCTTCCAAG 1823
QY 603 CTGCTTGGGTCTGTTTAGTTGGGTTGATCGTGGATTTTCTGCATGCAACCATAA 662
Db 1822 CTTCTCTCGCTTCTCCGCTGGGTTCATCGTGGATTTCTTGTGGAAGCGACGCTGA 1763
QY 663 TAGGGTTATGGAAGAGCAGCCAGGTGGTGTCTGCGAGCAACTAAATCGATCTTGT 722
Db 1762 CGGGATTTCATGGCGCGCGCGATCATGCTGCTGCGAGCAGCTCAAGGGATTGCTCG 1703
QY 723 GCCTTGAGCATTTACCCCATGGAGCTGATATCATATCAGTGATGCTCTGTTTCACCC 782
Db 1702 GGATCATCCATTTACGTGCGAGATGGGTTTCGTCCAGGTGATGACATCCGCTTCCAAG 1643
QY 783 AAACCTCATGAGTGGAGTGGGAAAGTGTGTGTAGGATGTGTCTTCAATTTCTCTCTCC 842
Db 1642 ACCACGAGGTGGGCGTGGCAGACATCTCATGGGCTCGCTTCTCTCGCGTCTCTCC 1583
QY 843 TTAGCACAGATACCTTCAGCAAAAAACGACCAAGGTTTTTTTTGGGTGTCAGCAATGGCGC 902
Db 1582 TCACCACACGCGCATCATGCGCCAGCAACCCAAAGCTTTCTGCGGTGTGAGCAGCTGCTC 1523
QY 903 CATTCAGCTCGTTATATTTGGGAAGTCTCTTGGTTATTTTCACTCACGCGGAGACGAG 962
Db 1522 CACTGACATCAGTGTATCTCTCATCATCTCGTTCGTTCAGCAAAAGC-----TCATG 1469
QY 963 GTGTTGAAGTATAGGAACTGGAAGAGGTTTGAATCCACCATCACTCACAAATCTCG 1022
Db 1468 GCATCAGTGTATTTGGCGATCTCCCAAGGATTTGAACCTCTCTCAGCGAATATGCTGA 1409
QY 1023 TATTTGTGTCGCTTTACATGACTACAGCTGTCAAACTGGCAATGTCTGTTGGCATCATAT 1082
Db 1408 CCTTCAGTGGCTCCTACGTAGGACTGGCCTTAAACACAGGAATATGACTGTCATACTAT 1349
QY 1083 CACTTGGGAGAGATACAGTAGGAAGAGCTTTTGCATGTATTAATAATTAACAATTTG 1142
Db 1348 CTCTCACTGAAGGATAGCAGTAGGAGGACATTTGCAATCCATCAACCACTACCAGTTG 1289
QY 1143 ATGGCAACAAAGAGATGATAGCTATTTGGGACCATGAAGCTAGTGTGTTCTTTTCACTCT 1202
Db 1288 ATGGAACAAGAGATGATGCGCATTTGGTGTATGAACATGCTGCTCATGTGCTCTCT 1229
QY 1203 GCTACTCTCAACACAGGACCATTTTCGCGTTCGGCTGTGAATCAATTAACCGTGGATGCAAGA 1262
Db 1228 GCTATGTCACACAGGATCGTTCTCGAGTTCGGCGGTGAACCTACAGCGCGGTGCAAGA 1169
QY 1263 CAGCAGCTTCCAAATTAATATGTCATCTTGAGTAATGTTGACATGTTATTTCTGACAC 1322
Db 1168 CGCGGCTGTGCAACATCGTGTGAGTGGGTGCGGCGGTGCTGAGTGTGTTCTGCTGATGC 1109
QY 1323 CCTTGTTCATTAACACTCCCTGCTGCTATAGCTATTAATCGTATCTGCAATGCTTTG 1382
Db 1108 CGCTGTTCACATACACCCCAACAGTGTCTCTCGCGGATCATCATCACCGCGGTGATCG 1049
QY 1383 GACTCATAGATTAAGAAGCAGCCATCTCATTTTAAAGTTTGACAAATTTGACTTTTGTG 1442
Db 1048 GCCTCATCGAGCTCCGGGCGCGCCAGGCTGTGGAGGTGCAAGCTCGACTTCTCTCG 989
QY 1443 TGTGATAGTGCATACATTTGGCTGGTCTTTGGCAGTGTGTAATTTGGCTTAGTCAATAG 1502
Db 988 CCGTCATGGCGCGCTTCTCGCGCTCTCTCTCGTCTCGGTCCAGATGGGCGCTCGCCATCG 929
```

QY	1503	CTATTGTAATATCTGTAACCTTCTGGGTACTTCTATTATTTGCAAGGCCAAGGACATTCGT	1562	Db	420	GCCTCGTCTGGCGCTGCACTACTTCTTCCCACTCTTCCAGTGGGGTCCGCTACAGCC	479
Db	928	CCGTGGCATCTCCCTCTTCAAGATCCTCTCAGGTCAACCGCCCAACATGGTGGTCA	869	QY	303	TTCAATTTTGAAGTCACTCATAGCTGGGATCACCATCGCTAGCTTGGCATTCCTC	362
QY	1563	TGGGCAACATCCAAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATGCAAAAC	1622	Db	480	CGCGCTCTGCGTCCGACCTCGTCGCGGCGCTCACCATTTGCCAGCCTCGCATCCGCG	539
Db	868	AGGGCGTGTGTCGGGCAAGCGAGCTACCGGAGCATGGCGAGTACAGGGAGGCCATGC	809	QY	363	AGGGCATCAGTTATGSCCAAGCTCGCAACCTCCCTCCAAATTTCTTGGACTATATTTCGAGCT	422
QY	1623	ATGTTCTCTGGAATGCTAAATTTCTAGAGATTGATGCACCAATTTACTTTTGCCAAATGGCAGCT	1682	Db	540	AGGGAATCAGCTACGCCAAGCTCGCCAACTTGCCTGCCCAANTCGTTGGCCTATATTCCAGCT	599
Db	808	GTGTGCGGTCTTCTCGTGGCGTCTGAGTCCGCACTACTTTCGCAACTCCATGT	749	QY	423	TTATACCAACCAATTTATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTG	482
QY	1683	ATTTAAGAGAAGGATCACAAGTGGATTGATGAAGAAGAAAGAAATTAAGTCTACAG	1742	Db	600	TCGTGCGCGCTCATCTACGGCTGTCTGGGAGCTCGCGGACCTGCGGGTGGGCGCG	659
Db	748	ACCTCGCGAGAGGATCATGAGTTTCTCCGGAGGAAGACAGCGGGCGGCAAGTGCA	689	QY	483	TGGCGTTTGGATTCGTATGCGTTCTGATGGGTTCGATGTTGAGTAAATGCCGCTTGAATCCCAATGAAG	542
QY	1743	GGGAGACTAGTTTGCAGTATGTATAATTTGATATGAGTCTGTTGGAACATTTGATACAA	1802	Db	660	TGTCATCGCGTGTGTTGATGGGTCCATGCTCCGGAGCGCGGTGTCGCGGACGAGC	719
Db	688	ACCAGTGCCTGTTCAGATGCATCATCTCGCATGAGTGTCTTGCAGCGATCGACACGA	629	QY	543	ACCCAAAGCTTTTACCTCCACCTGGCTTTCACAGCTACATTAATTTGCTGCTGTTTTCAGG	602
QY	1803	GTGGAAATAGTATGCTTTGAAGAGGTCAAGAAGATTACAGAGAGAAGAGCTTACAGCTTG	1862	Db	720	AGCGCTCTCTTACCTGACGTGGCTTTCACCGCCACCTCTTTCGCCGCGCTCTTCCAGG	779
Db	628	GTGGCTCGATGCATTTGCAGAGCTGAAGAGTGTCTGGAAAAGAAACATCGAGCTTG	569	QY	603	CTGCCCTTGGGTCTGTTTGGGTTCGATCGTGGATTTTCTGTCAATGCAACATATAA	662
QY	1863	TTTTGTCAATCTGTAAAGTGAAGTGAAGAACTGAACAAATCGAAATGTTCCAAATC	1922	Db	780	CGTCCCTGGGATTCCTCAGGCTGGGCTTTCATCGTGGACTTCTGTCCAAAGCGACGCTGA	839
Db	568	TGTCGCCCAACCGGTGGATCGGTGACCGAGAGGCTGTACAACTCGTGGTTCGGCAAGA	509	QY	663	TAGGTTTCAATGGGAGGAGCAGCGGTGGTGTCTGCGAGCAACTAAAAATGATCTTGT	722
QY	1923	ATTTAGGAGAAGATGATCTATCTGACTGTTTGAAGAGCGGTTGGAGC	1971	Db	840	CGGGCTTCATGGCGCGCGCGCTCATCGTGTCTGTCGACGAGCTCAAGGGCTCTGCTCG	899
Db	508	CGTTCGGCTCAGACCGCGTGTCTTCAGCGTCCGAGCGCGTTCGCGGC	460	QY	723	GCCTTGAGCATTTACCCATGGAGCTGATATCATATCATAGTATGCTGCTGTTTTCACCC	782
RESULT 13				Db	900	GCATCTCCACTTACCTCCACATGGGATTCCTCGACGTATGCGCTCCGCTGCTCAACC	959
US-10-425-114-15719				QY	783	AAACTCATGAGTGGAGTGGGAAAGTGTGTGTAGGATGTGTCTTCAATTTTCTTCTCC	842
; Sequence 15719, Application US/10425114				Db	960	GCCACGACGATGGGAAGTGGCAGACGATCGTCATGGGCTCCGCTTCTCGCCATCTCTC	1019
; Publication No. US20040034888A1				QY	843	TTAGCACAAGATATCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGC	902
; GENERAL INFORMATION:				Db	1020	TCCTCAGCGGCCAAATCAGCGCCAGGAAACCCAAAGCTTTTCTGGGTATCAGCAGGTGCTC	1079
; APPLICANT: Liu, Jingdong				QY	903	CATTGACGTCGGTTATATTTGGGAAGTCTCTTGGTTTATTTTCACTCACCGCGAGAGCAG	962
; APPLICANT: Zhou, Yihua				Db	1080	CCCTGGCGTGGTGTATCTCCACCTCTCTCTCTTCACTGGAATATCCCCAGCATCA	1139
; APPLICANT: Kovalic, David K.				QY	963	GTGTTGAAGTGTATAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCAAAATCTGG	1022
; APPLICANT: Screen, Steven E.				Db	1140	GTGTT-----ATTGGCATCTCCCGAGGGAGTGAACTTCTTGGGGAACATGCTCA	1193
; APPLICANT: Tabaska, Jack E.				QY	1023	TATTTGTGTCGCTTATACATGACTACAGCTGTCAAAACTGGGCAATGCTGTTGGCATCATAT	1082
; APPLICANT: Cao, Yongwei				Db	1194	GCTTCAGCGGCTCTTATGTGGCGTGTGAGATCAAAACCGGGATCATGACAGGCATCTGT	1253
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With				QY	1083	CACTTTGGCGGAGGATAGCAGTAGGAAGAGCTTGTGCAATGTATATAAATTTACAATATG	1142
; FILE REFERENCE: 38-21(53313)B				Db	1254	CCTTAACGAAGAGGATCGCAGTGGGACAGACCTTCGCGTCCATCAACAACTACCAAGTGG	1313
; CURRENT APPLICATION NUMBER: US/10/425,114				QY	1143	ATGGCAACAAAGAGATGATAGCTATTTGGGACCATGAACGTAGTTGGTTCTTTCACTCTT	1202
; CURRENT FILING DATE: 2003-04-28				Db	1314	ACGGAAACAGGAGATGATGGGATCGGGCTGATGAACATGCGGGCTCTTCCGCTCTCT	1373
; NUMBER OF SEQ ID NOS: 73128				QY	1203	GCTACTCACAACAGGACCAATTTTCGCTTGGGTGTGAACTATAACCGTGTGATCAAGA	1262
; SEQ ID NO 15719				Db	1374	GCTACGTGACGACGGGCTCTTCTCCCGTGGCGGTGAACCTTACAGCGGGCTGACAGA	1433
; LENGTH: 2656				QY	1263	CAGCAGCTTCCAACTTATATGCTTGCAGTAACTGTTCAGATTTGTTGACATTTTATTCGACAC	1322
; TYPE: DNA				Db	1434	CGGCGCTGTCCAACTCGTGTGATGGCGGCGGGTCTGGTGAACGCTGTCTTCTTCATGC	1493
; ORGANISM: Zea mays				QY	1323	CTTGTGTTCAATFACATCTCCCTGGTGGTGTATCAGCTATTATCGTATCTGCAATGCTTG	1382
; FEATURE:				Db	1494	CGCTGTTTCACTACACCGGACGATGATCTTCGGCGCGATCATCATCAGCGGGTGGTGG	1553
; OTHER INFORMATION: Clone ID: LIB3060-059-F10_FLI							
US-10-425-114-15719							
Query Match		23.2%;	Score 567;	DB 18;	Length 2656;		
Best Local Similarity		57.0%;	Pred. No. 2.1e-141;				
Matches 1060;	Conservative	0;	Mismatches 795;	Indels	6;	Gaps	1;
QY	123	TGCACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTTGA	182	Db	300	TGCACAGGTGTCGTGCGGAGCGGGGAGCAGCGGAGGCGTTCGGCGCGGCTGG	359
Db	183	AGGAGACTTCTTCTCCCTGATGACCCCTTTGAGGCGATTCAAGAACAGCGAGCTTCCAAAG	242	QY	360	CCGAGGTGTTCTTCCCGGACGACCCGCTGCACAGTTCAAGAACCAAGTGTGCGCGCGGC	419
Db	243	AGTTCAATGCTTGGCTTCACTTCTTCCCAATTTTCGATGGGCTCCCAATACACTT	302	QY			

```
QY 1383 GACTCATAGATTATGAAGCAGCCATCATCTATTTAAAGTTGACAAATTTGACTTTGTGG 1442
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1554 GGTGGTGGACGTGCGCGGCGCCAGGCTGTGGAGGTGGACAGCTGGACTTCTCTGG 1613
QY 1443 TGTGCATAGTGCATACATTGGCGTGGCTTTGGCAGTGTGAAATTTGGCTTAGTCATAG 1502
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1614 CGTGGTGGCGGCTTCTCGGCGTGTCTGTGTCTCGTGGCAGACGGGCTGGGCGTGG 1673
QY 1503 CTATTGTATATCTGTACTTCGGGTACTTCTATTATTCGAAGGCCAAGACATTCGTTT 1562
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1674 CGTCCGGCATCTCGCTCTTCAAGGTCTGTGAGGTCAACCGCCCAAGCTGTGGTGG 1733
QY 1563 TGGGCAACATTCCAAATCTGTGATATACCGAAATCTTGAGCACTATCAAAATGCAAAAC 1622
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1734 AGGCGCTCGTCCCGGCGGACGAGAGCTACCGAGCTGGCGCAGTACCGCGGCGCTCC 1793
QY 1623 ATGTTCTTGGAAATGCTTAATTTCTAGAGATTGATGACCAATTTACTTTTGGCAATGCGAGCT 1682
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1794 GCGTGGCGGCTTCTCTCGTGGTGTGAGTCCGCGCTACTTTCGCCAACTCCATGT 1853
QY 1683 ATTTAAGAGAAAGGATCAAGGTGATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1742
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1854 ACCTGTGGAGCGGCTCATCGCTACTCTCGCGACGAGGAGGAGCGCGCTCAAGTCCA 1913
QY 1743 GGGAGACTAGTTTGCAGTATGTTATAATTGATGATGCTGTGTGAAACATTTGATACAA 1802
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1914 ACCACCTTCCATCCGATGGTCTCTGACATGGCGCGCTGCGGCGATGACACAGA 1973
QY 1803 GTGGAATAAGTATGCTTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGTACAGCTTG 1862
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1974 GCGGTCTAGACGGCTGTCGAGCTCAAGAAAGTCTGTGACAAAAGAAACATCGAGCTGG 2033
QY 1863 TTTTGTCAATCTGTAAAGTGAAGTATGAAGAACTGAACAAATCGAAGTTCGAAATC 1922
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2034 TGTTCGCAACCGGTGGGTCGTCGCGAGAGAGATGTTCAACTCGGCGGTGGCGAGA 2093
QY 1923 ATTTAGGAGAAATGATCTATCTGACTGTTGAAGAGCGCGTGGAGCATGCAACTTCA 1982
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2094 GCTTCGGTGGCGCGCTCTTCTTTCAGGTAGCGAGAGCGCGTGGCGGCGGTGCA 2153
QY 1983 A 1983
DB 2154 A 2154
```

RESULT 14

```
US-10-425-115-75763
; Sequence 75763, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 75763
; LENGTH: 2713
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_169123C.1
US-10-425-115-75763
```

Query Match 23.2%; Score 567; DB 20; Length 2713;
Best Local Similarity 57.0%; Pred. No. 2.1e-141;
Matches 1060; Conservative 0; Mismatches 795; Indels 6; Gaps 1;

```
QY 123 TGCACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTTCAAGTCTCTTAAAGTACTCTTTGA 182
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
357 TGCACCAAGTGTGCTGCGGAGCGCGGACGACGCGAAGCGCTGCGGAGCGCTGG 416
QY 183 AGGAGACTTTTCTTCCCTGATGACCCCTTTGAGGCGAGTTCAAGAACAGCAGCTTCCAAGA 242
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
417 CCGAGGTGTTTCTTCCCGGACGACCGCTGACACAGTTCAAGAACAGTCTGTCGGCGCGC 476
QY 243 AGTTCAATGCTTGGCCTTCAGTTCTTCTTCCCATTTTCCAAATGGGCTCCCAAATACACCT 302
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
477 GCCTGCTGCTGGGCTGCACTACTTCTTCCCATCTTCCAGTGGGGTCCGCTACAGCC 536
QY 303 TTCAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATGCTAGCTTTGGCATTCTTC 362
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
537 CGCGCTCTCTGGCTCCGACCTGTCGCGCGCTCACCAATGCCAGCTTCGCAATCCGCG 596
QY 363 AGGCGCTCAGTTATGCGAGCTCGCCAACTCCCTCCAATTTCTTGGATATATATTCGAGCT 422
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
597 AGGGAATCAGCTACGCCAAGCTCGCAACCTGCGGCCAATCGTTGGCTTATATTCAGCT 656
QY 423 TTATACCAACCATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACTG 482
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
657 TCGTCCCGCGCTCATCTACGCGCTGCTGGGAGCTCGGGGACCTGGGGTGGGCGCG 716
QY 483 TGGCGGTTGGATCGCTTCTGATGGGTTGATGTTGAGTAATGCCGCTTGAATCCAAATGAAG 542
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
717 TGTCCATCGCTGCTGCTGATGGGTCCATGCTCCGGGACGCCGTGCGCGGACGAGC 776
QY 543 ACCCAAAGCTTTACTCAGCTGGCTTTCAAGCTACATATTTGCTGCTGTTTTCAGG 602
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
777 AGCCGCTCTCTACCTGCACTGGCTTTCACCGCCACCTTCTTCCCGCGCTTTCAGG 836
QY 603 CTGCTCTGCTGCTGTTAGTTGGGTTGATCGTGGATTTTCTGTCACATGCAACATATAA 662
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
837 GCTCTCTGGATTCCTCAGCTGGCTTCATCGTGAGCTTCTGTCGAAGGCGAGCTGA 896
QY 663 TAGGCTTCATGGGAGGAGCAGCAGCTGGTGTGTCGACGAACTAAATCGATCTTTG 722
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
897 CGGGCTTCATGGCGGCGCGCGCTCATGCTGCTGCGCAGCTCAAGGCGCTGCTCG 956
QY 723 GCCTTGAGATTTCAACCATGGAGCTGATATCATATGATGCTGCTGTTTTCACCC 782
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
957 GCATCTCCCACTTCACTCCCACTGCGGATTCCTGACGTCATGCGCTCCGCTCGTCAACC 1016
QY 783 AAACCTCATGAGTGGAGTGGGAAAGTGTGTTAGGATGTGCTTCTCATTTTCTTCTCC 842
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1017 GCCACGAGTGGAGTGGCAGCATGCTGATGCGGCTCCGCTTCTCGCCATCTCTC 1076
QY 843 TTAGCACAGATACCTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGC 902
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1077 TCCTCACGCGCCAAATCAGCGCCAGGAACCCAAAGCTTTTCTGGGTATCAGCAGGTGCTC 1136
QY 903 CATTCAGCTCGGTTATATTTGGGAACTCTCTGGTTTATTTACTCAGCGCGGAGAGCAGG 962
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1137 CCTGGCGTGGTGTATCATCTCCACCATCTCTCTCTCATCTGGAATCCCCAGCATCA 1196
QY 963 GTGTGAGTGTAGAGAACTGAAGAGGGTTGAATCCCACTCACTCAAAATCTGG 1022
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1197 GTGTT-----ATTGGCATCTCCCGGAGAGTGAACCCCTCTTCGGGAAACATGCTCA 1250
QY 1023 TATTTGTGTGCGCTTACATGACTACAGCTGTCAAACTGGCATTTGCTGTGGCATCATAT 1082
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1251 GCTTCAGCGGCTCTTATGTGGCGCTGACGATCAAAACCGGGATCATGACAGGCATCTGT 1310
QY 1083 CACTTGGGAAAGGATAGCAGTAGGAGAACCTTTGCAATGTATATAAATTTACAATATTTG 1142
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1311 CCTTAACAGAGGGATCGCAGTGGGAGGACCTTCGCGTCCATCAACAACATACAGGTGG 1370
QY 1143 ATGGCAACAAAGAGATGATGCTATTTGGGACCATGAACTGATGTTTCTTTCACCTCTT 1202
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1371 ACGGAAACAAAGAGATGATGCGGATCGGGCTGATGAACATGCGGGCTCTCTCGGCTCTC 1430
QY 1203 GCTACTCTACACAGGACCAATTTTCGCGTTCGGCTGTGAACATATAACGCTGGATGCAAGA 1262
```

[illegible]

RESULT 15
 US-10-424-599-32308
 ; Sequence 32308, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 32308
 ; LENGTH: 536

```

; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129179C.1
US-10-424-599-32308

Query Match          21.2%; Score 520.2; DB 18; Length 536;
Best Local Similarity 98.5%; Pred. No. 3.3e-129;
Matches 525; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 7 GCTAGCTCGCACATTAAGTTTATATAACACATATTTGCTGCTTTAGAAATACTATTATGA 66
   |||||
Db 4 GCTAGCTCGCACATTAAGTTTATATAACACATATTTGCTGCTTTAGAAATACTATTATGA 63

Qy 67 AGATATGGGAGTGTTAGATTATGAGTACCCCTTTGGGCATGAAACAACCTTTGAGAGATGCA 126
   |||||
Db 64 AGATATGGGAGTGTTAGATTATGAGTACCCCTTTGGGCATGAAACAACCTTTGAGAGATGCA 123

Qy 127 CCAAGTCGAGGTTCCACCGGCACACGCGCTTTTTCAAAGCTCTTAAAGTACTCTTTGGAAGGA 186
   |||||
Db 124 CCAAGTCGAGGTTCCACCGGCACACGCGCTTTTTCAAAGCTCTTAAAGTACTCTTTGGAAGGA 183

Qy 187 GACTTTCTTCCCTGATGACCTTTTGAGGCAAGTTTCAAGAAACAAGCGAGCTTCCAAAGATT 246
   |||||
Db 184 GACTTTCTTCCCTGATGACCTTTTGAGGCAAGTTTCAAGAAACAAGCGAGCTTCCAAAGATT 243

Qy 247 CATGCTTGGCCCTTCAGTTCTTCTTCCCAATTTTCGAATGGGCTCCCAAAATACACCTTTCA 306
   |||||
Db 244 CATGCTTGGCCCTTCAGTTCTTCTTCCCAATTTTCGAATGGGCTCCCAAAATACACCTTTCA 303

Qy 307 GTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGCGCCATTCCTCAGGG 366
   |||||
Db 304 GTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGCGCCATTCCTCAGGG 363

Qy 367 CATCAGTTATGCCAAAGCTCGGCCAACTCCCTCCAAATTCCTTGGACTATATTCGAGCTTTAT 426
   |||||
Db 364 CATCAGTTATGCCAAAGCTCGGCCAACTCCCTCCAAATTCCTTGGACTATATTCGAGCTTTAT 423

Qy 427 ACCACCAATTCGATTTATGCGATGATGGGTAGCTCGAGGGAATTCGGCAGTGGGGACTGTGGC 486
   |||||
Db 424 ACCACCAATTCGATTTATGCGATGATGGGTAGCTCGAGGGAATTCGGCAGTGGGGACTGTGGC 483

Qy 487 GGTGTGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTGATCCCAATG 539
   |||||
Db 484 GGTGTGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTGATCCCAATG 536

```

Search completed: September 12, 2005, 14:59:41
Job time : 5645 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2005, 11:07:23 ; Search time 57 Seconds
(without alignments)
890.550 Million cell updates/sec

Title: US-10-762-049-18
Perfect score: 3475
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGWNV 680

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3475	100.0	680	4	US-09-720-317A-18
2	2642.5	76.0	658	4	US-09-720-317A-24
3	2575.5	74.1	646	4	US-09-720-317A-26
4	2314	66.6	646	4	US-09-720-317A-28
5	2164	62.3	579	4	US-09-720-317A-4
6	1901	54.7	621	4	US-09-720-317A-16
7	1890.5	54.4	631	4	US-09-720-317A-29
8	1843.5	53.1	688	4	US-09-720-317A-2
9	1798.5	51.8	656	4	US-09-720-317A-20
10	1760.5	50.7	660	4	US-09-720-317A-30
11	1723.5	49.6	660	4	US-09-720-317A-23
12	1503	43.3	644	4	US-09-720-317A-25
13	997	28.7	685	4	US-09-720-317A-31
14	961.5	27.7	593	4	US-09-720-317A-22
15	731.5	21.1	226	4	US-09-720-317A-8
16	687.5	19.8	233	4	US-09-720-317A-27
17	687.5	19.8	536	4	US-09-252-991A-23812
18	675.5	19.4	311	4	US-09-720-317A-6
19	650.5	18.7	744	4	US-09-785-381-3
20	650	18.7	744	4	US-09-785-381-1
21	643.5	18.5	590	4	US-09-902-540-14944
22	641.5	18.5	748	4	US-09-949-016-10387
23	612	17.6	565	4	US-09-602-787A-616
24	587.5	16.9	780	4	US-09-785-381-11
25	586.5	16.9	764	1	US-08-424-567-2
26	586.5	16.9	764	2	US-08-711-928-2
27	586.5	16.9	764	3	US-09-184-937-2

28	586.5	16.9	790	4	US-09-949-016-11220	Sequence 11220, A
29	576.5	16.6	803	4	US-09-949-016-11498	Sequence 11498, A
30	567	16.3	828	4	US-09-248-796A-20746	Sequence 20746, A
31	549.5	15.8	663	4	US-09-875-811-6	Sequence 6, Appli
32	549	15.8	656	4	US-09-875-811-10	Sequence 10, Appl
33	546.5	15.7	679	4	US-09-875-811-2	Sequence 2, Appli
34	489	14.1	616	4	US-09-543-681A-4421	Sequence 4421, Ap
35	485.5	14.0	605	4	US-09-875-811-8	Sequence 8, Appli
36	485	14.0	598	4	US-09-875-811-12	Sequence 12, Appl
37	482.5	13.9	621	4	US-09-875-811-4	Sequence 4, Appli
38	454	13.1	970	4	US-09-795-927-7	Sequence 7, Appli
39	449	12.9	535	4	US-09-252-991A-21805	Sequence 21805, A
40	426.5	12.3	472	4	US-09-902-540-13374	Sequence 13374, A
41	402.5	11.6	575	4	US-09-438-185A-1015	Sequence 1015, Ap
42	381	11.0	159	4	US-09-720-317A-14	Sequence 14, Appl
43	369	10.6	566	4	US-09-543-681A-4544	Sequence 4544, Ap
44	363	10.4	533	4	US-09-107-532A-5244	Sequence 5244, Ap
45	360.5	10.4	147	4	US-09-720-317A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-720-317A-18
; Sequence 18, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Glycine max
US-09-720-317A-18

Query Match	100.0%;	Score	3475;	DB	4;	Length	680;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	680;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0;
Qy	1	HELARTLSYTHICLLRNTI	IEDMGSDVDEYPLGMNPFVHVEVPPQPFKSLKYS	60			
Db	1	HELARTLSYTHICLLRNTI	IEDMGSDVDEYPLGMNPFVHVEVPPQPFKSLKYS	60			
Qy	61	KETFPDDPLRQKPKASKFMLGLQFPFPIFEWAPKYTFQFLKADLIAGITIASLAIP	120				
Db	61	KETFPDDPLRQKPKASKFMLGLQFPFPIFEWAPKYTFQFLKADLIAGITIASLAIP	120				
Qy	121	QGISYAKLANLPPIILGLYSSFTPLIYAMGSSRDLAGTAVGSLMGSMLSNVDPNE	180				
Db	121	QGISYAKLANLPPIILGLYSSFTPLIYAMGSSRDLAGTAVGSLMGSMLSNVDPNE	180				
Qy	181	DKPLYHLAFTATLPAQVFOALGLIFRLGLIVDFLSHATIIIGFMGSAATVVCQLQKLSIL	240				
Db	181	DKPLYHLAFTATLPAQVFOALGLIFRLGLIVDFLSHATIIIGFMGSAATVVCQLQKLSIL	240				
Qy	241	GLEHETHGADIISVMRSVFTQTHWRWSAVLGCVFIFFLSTRYFSKCRPRFFWVSAMA	300				
Db	241	GLEHETHGADIISVMRSVFTQTHWRWSAVLGCVFIFFLSTRYFSKCRPRFFWVSAMA	300				
Qy	301	PLTSTVLGSLLYFTHAEKHGVEVIGELKGLNPPSLTNLNVFVSPYMTTAVKTGIVGII	360				
Db	301	PLTSTVLGSLLYFTHAEKHGVEVIGELKGLNPPSLTNLNVFVSPYMTTAVKTGIVGII	360				

RESULT 12
US-09-720-317A-25
; Sequence 25, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Stylosanthes hamata
US-09-720-317A-25

Query Match 43.3%; Score 1503; DB 4; Length 644;
Best Local Similarity 47.0%; Pred. No. 1.2e-142;
Matches 301; Conservative 124; Mismatches 193; Indels 22; Gaps 6;

QY 33 LGNNFERVHQ--VEVPPQPFKSLKYSIKETFFDDPLRQFNKPKASKKMLGQF-- 88
DB 4 LGTEQSESRQWLNPNPPLTKFLGLKD-----NKFSTSSSKKTRAVSFLA 55
QY 89 -FPPIEWAPKYTFQFKADLIAGITIASIAIPOGISYAKLANLPILGLYSFIPPLIY 147
DB 56 SLFPILSWRTYSATFEKDDLLSGLTASISIQISGYANLAKLDPOYGLYSVIPVIY 115
QY 148 AMGSSRDLAGVTAVGSLMGLMNAVDNPDPKLYLHLAFTATLFAGVFOAALGLPR 207
DB 116 ALMGSSREIAIGPVAVVMSLSSLPVKVIDPDADPNDRNLVPTVTLFAGIFQAFGLR 175
QY 208 LGLVDFLSHATIIIGWGNATVVCLOOLKSLIIGLEHFTGADIISVMSVFTQTHE 264
DB 176 LGLVDFLSHALVGMAGAAIIVIGLQKGLGLTHFTTKTDAVALKSVYTSLHQIT 235
QY 265 ---WRWESAVLCGVFIFFLLSRYFSKKRPRFVWSAMAPLTSVILGSLVYVTHAEKH 320
DB 236 SSENWSPFNVICCSFLIFLAARFIGNRNKKEFNLPAIAPLLSVILSLVFLSKGDH 295
QY 321 GVEVIGELKGLNPPSLTNLTVFVSPYMTTAVKTGIVVGIISLAEGIAVGRSPFAMYN 380
DB 296 GVNIKHVQGLNPSSVHKQLQNGPHVGOAAKIGLISAIITAEIAVGRSPFANIKGYHL 355
QY 381 DGNKEMIAIGTMNVGSGFTSCYLTGTPFSRSVAVNAGCKTAASNIIMSLAVMLTLLPLT 440
DB 356 DGNKEMLAGCMGNIAIGLSCTSYSTGTSFRTAVNFSAGCKTAVSNVMAVTLLELFT 415
QY 441 PLPHYTPLVLSAIIYSAMGLIDYEAHLFKVDKDFVFCMSAVIGVVFVSGVETGLVI 500
DB 416 RLITYTPMAILASIIISALPGLIDIGEAHYIWKVDKDFLACLFAGFVLVFSIELGLI 475
QY 501 AIVISVLRVLLFTFRTFVLGNIPNSVIRYRNVEHYQNAKHVPGMLILBIDA-PIYFANA 559
DB 476 ALSISPAKILLQAIRPGVEVLGRIPTEAYCDVAQPMVAVTPGLIVIRISSGLCFANA 535
QY 560 SYLRERITRWI-DEBERIKATGETSLOVVIDMSAVGNIDTSGISMLSEVKKITERREL 618
DB 536 GFVREIRLKWVEDEEQNIEEAAKGRVQAIIDMTDLTNVDTSGILALELHKLLSRGV 595
QY 619 QLVNVPVSVMKLNKSKFQNLGKKWLYLTVEEAVGAC 658
DB 596 ELAWNPRNEVHKLKVFVDKIGKERVFLTVAEAVDAC 635

RESULT 13

US-09-720-317A-31
; Sequence 31, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-720-317A-31

Query Match 28.7%; Score 997; DB 4; Length 685;
Best Local Similarity 34.0%; Pred. No. 1.8e-91;
Matches 217; Conservative 140; Mismatches 232; Indels 50; Gaps 13;

QY 47 PPPQPFKSLKYSIKETFFDDPLRQFNKPKASKKMLGLOFFFPPIFEWAPKYTF-QFLK 105
DB 51 PSISFP-----DIFSGWTAKIKRMRLVDWIDTLFPCFRVIRTYRWSEYFK 96
QY 106 ADLIAGITIASIAIPOGISYAKLANLPILGLYSFIPPLIYAMGSSRDLAGVTAVGS 165
DB 97 LDLMAGITVGMILVPMQMSYAKLAGLPIYGLYSFVPVVFVAIFGSSRLAIGPVALVS 156
QY 166 LLMGSMLSNVDNPDPKLYLHLAFTATLFAGVFOAALGLFRLGLVDFLSHATIIIGFMG 225
DB 157 LLVSNALGGIADTNBE--LHIELAILLALLVILECIMGLLGLMLIRFISHVSGFTS 214
QY 226 GAATVVCLOQLKSLIIGLEHFTGADIISVMSVFTQTHEWRWESAVLCGVFIFFLLSTRY 285
DB 215 ASAVIGLSQIKYFLGYS-IARSSKIVPIVESIAGADKQFPWPFVMSGLILVILQMKH 273
QY 286 FSKKRPFRFWSAMAPLTSVILGSLVYVTHAEKHGVEVIGELKGLN----PPSLTNLV 341
DB 274 VGKAKKELQFLRAAAPLGTGIVLGTITAKVFHPP--SISLVGEIPQGLPTFSFPRSPDHAK 331
QY 342 FVSPYMTTAVKTGIVVGIISLAEGIAVGRSPFAMYNKYNIDGNKEMIAICTMNVGSGFTSC 401
DB 332 TLLP--TSALITG-----VPILSVGIAKALAAKNRYELDSNLSLFGLVANILGSLFSA 384
QY 402 YLTGTPFSRSVAVNAGCKTAASNIIMSLAVMLTLLPLTFLPHYTPLVLSAIIYSAMLG 461
DB 385 YPATGSFERSAVNNSEAKTGLSGLITGIIIGCSLLFTPMFKYIPOCALAAIIVISAVSG 444
QY 462 LIDYEAHLFKVDKDFV--CMSAVIGVVFVSGVSEIGLVIAIVISVLRVLLFIARPRTFV 520
DB 445 LDVYDEAIFLWRVDRKDFSLWTITSTITLFFG-IEIGVLVGVGFSLAFVHESANPHIAV 503
QY 521 LGNIPNSVIRYRNVEHYQNAKHVPGMLILBIDAPIYFANASYLRERITRW---IDE----- 572
DB 504 LGRLPGTTVYRNKIQPEAYTYNGVIVRIDSPIFYANISYIKDRUREVEAVADKTYNRG 563
QY 573 -DEBERIKATGETSLOVVIDMSAVGNIDTSGISMLSEVKKITERRELQVLVNPVSEVMK 631
DB 564 LEVDRI-----NFVILEMSPVTHIDSSAVAEALKELYQYKTRDIQIAINPNKDVHL 615
QY 632 KLNKSKFQNLGKKWLYLTVEEAVGACNPNLRAKNTPK 670
DB 616 TIARSGMVELVGKWFVRVHDVAVQVCLQYVQSSNLEDK 654

RESULT 14

US-09-720-317A-22

; Sequence 22, Application US/09720317A

```
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-720-317A-22

Query Match      27.7%; Score 961.5; DB 4; Length 593;
Best Local Similarity 35.0%; Pred. No. 5.5e-88;
Matches 204; Conservative 127; Mismatches 235; Indels 17; Gaps 7;

QY  91 PIFEWAPKTF-QFLKADLIAGTIIASLAIPQGISYAKLANLPPTILGLYSSFTIPPLIYAM 149
Db  22 PCLAMNRSYRWKEDFOADLAAGITVGMLVPQAMSYAKLAGLHPHYGLTYGVPLFVYAI 81

QY  150 MGSSRDLAVGTVAVGSLLMGMSLNAVDNEDPKLYLHLAFTATTLFAGVFOALGIFRIG 209
Db  82 FGSSRLQAGVALVSNLVGGIV--NSSSELYTELAIIAFLAFVVGILECLMALLRIG 139

QY  210 LIVDFLUSHATIIGFMGGAATVVCLOQLKSIILGLEHTHGADIISVNRVFTQTHEWRWES 269
Db  140 WLIRFTSHSVISGFTTASAIIVIGLSQIKYFLGYS-VTRSSKIIPLETIESIAGIDQFSWPP 198

QY  270 AVLGCVFIFPLSTRVFSKRPFRFWVSAMAPLTSVILGSLAVYFTHAEKHGVVEIGELK 329
Db  199 FWGSAFLVILLIMKLGKTNKRLFRASGPLTAVVLGTFLFKIFRPT--AISVVGEIP 256

QY  330 KGLNPPLTNLVFVSP-----YMTAVKTGIVVGIISLAEGIAVGRSFAMYNKYNIDGNKE 385
Db  257 QGL--PS-----PSIPRGFEHLMSLMPAILITGVVAILESVGIKALAAKNGYELDSNKE 309

QY  386 MTAIGTMNVVGSFTSCYLITGPFPSRSVAVNAGCKTAASNIIMSLAVMLTLLPLTPLPHY 445
Db  310 LFLGLSLNICGSPFSAYPATGSPFSRSVAVNHESAKTGLSGIIMGIIICSALEFMTPLFTD 369

QY  446 TPLVLSAIIVSAMGLIDYEAHLFKVDKDFVVCMSAYIGVVGFSVEIGLVIATVIS 505
Db  370 IPQCALAAIVISAVTGLVDYEEAIFLWGDIDKDDFLWAMTFTTTLFFGIEIGVLGVGFS 429

QY  506 VLRLVLLFIARPTFVLGNIPNSVIYRNVHYQNAKHVPGLMLILEIDAPIYFANASYLRER 565
Db  430 LAFVHESANPHAVLGRLPFGTIVYNTLOYEAYTYNGIVVVRVDADAPIYFANISYIKDR 489

QY  566 ITRWIDEEBERIKATGETSLQVVIIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNP 625
Db  490 LREYELKLPNSNRGPDVGRVYFVILEMSPVTVIDSSAVQALKDLHQEYKARDIQIAIANP 549

QY  626 VSEVMKKLNKSKFQNHGKKWLYLTVEEAVGACNENLRASKTN 668
Db  550 NRQVHLLLSRAGIIDMIGAWCFVRVVDHAVQVCLQHVSRSSSN 592
```

```
RESULT 15
US-09-720-317A-8
; Sequence 8, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
```

```
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helianthus tuberosus
US-09-720-317A-8

Query Match      21.1%; Score 731.5; DB 4; Length 226;
Best Local Similarity 65.6%; Pred. No. 2e-65;
Matches 145; Conservative 31; Mismatches 38; Indels 7; Gaps 1;

QY  467 RAIHLFKVDKDFVVCMSAYIGVVGFSVEIGLVIATVISLRLVLLFIARPTFVLGNIPN 526
Db  6 AAIHLLWTLDKDFVVCMSAYIGVVGFSVEIGLVIATVALSLKRLVLLFVSRPRTSTLGLIPD 65

QY  527 SVIYRNVHYQNAKHVPGLMLILEIDAPIYFANASYLRERITRWIDEEBERIKATGETSIQ 586
Db  66 STIYRMDQYQNAKSVPGILLIQIEAPIYFANSSYLRRERIVRWVDEEDRLKSLKENDLQ 125

QY  587 VYIIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKKLNKSKFQNHGKKW 646
Db  126 VYILALSAVGNIDTSGITMLGEVKKWMERRGLKVLANPGGSEVIKKMNKAKLIEVIGQEW 185

QY  647 IYLTVEEAVGACNENLRASKTNPKK-----DETEGWNV 680
Db  186 IYLTVEEAVGACNFMHLHYKNAEKPTSGSESGKSRNDNV 226
```

Search completed: September 12, 2005, 18:43:11
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2005, 18:35:01 ; Search time 1408 seconds
(without alignments)
190.492 Million cell updates/sec

Title: US-10-762-049-18
Perfect score: 3475
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGWNV 680

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1777461 seqs, 394431504 residues

Total number of hits satisfying chosen parameters: 1777461

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pap.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pap.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pap.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pap.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pap.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2330.5	67.1	657	US-10-437-963-140667	Sequence 140667,
2	2015	58.0	414	US-10-424-599-273958	Sequence 273958,
3	1853.5	53.3	695	US-10-425-115-235264	Sequence 235264,
4	1850	53.2	659	US-10-424-599-207945	Sequence 207945,
5	1811.5	52.1	666	US-10-437-963-198336	Sequence 198336,
6	1773.5	51.0	662	US-10-437-963-179488	Sequence 179488,
7	1762	50.7	681	US-10-425-115-260426	Sequence 260426,
8	1733.5	49.9	656	US-10-437-963-175893	Sequence 175893,
9	1724	49.6	397	US-10-424-599-273959	Sequence 273959,
10	1723.5	49.6	653	US-10-437-963-179495	Sequence 179495,
11	1722.5	49.6	662	US-10-425-115-366870	Sequence 366870,

12	1563.5	45.0	659	16	US-10-437-963-174545	Sequence 174545,
13	1500	43.2	653	15	US-10-424-599-195017	Sequence 195017,
14	1460	42.0	575	15	US-10-425-114-69272	Sequence 69272, A
15	1436	41.3	405	16	US-10-425-115-228851	Sequence 228851,
16	1406	40.5	640	16	US-10-425-115-224859	Sequence 224859,
17	1379.5	39.7	687	15	US-10-424-599-217235	Sequence 217235,
18	1344.5	38.7	587	16	US-10-437-963-180840	Sequence 180840,
19	1222	35.2	490	15	US-10-424-599-225979	Sequence 225979,
20	1197.5	34.5	730	16	US-10-437-963-178936	Sequence 178936,
21	1185.5	34.1	543	15	US-10-424-599-217231	Sequence 217231,
22	1018.5	29.3	390	15	US-10-424-599-282576	Sequence 282576,
23	1013	29.2	392	15	US-10-424-599-197192	Sequence 197192,
24	1011.5	29.1	339	15	US-10-424-599-273793	Sequence 273793,
25	998	28.7	485	16	US-10-739-930-9381	Sequence 9381, Ap
26	994	28.6	485	15	US-10-424-599-203123	Sequence 203123,
27	975.5	28.1	392	16	US-10-425-115-260428	Sequence 260428,
28	955.5	27.5	689	16	US-10-437-963-195183	Sequence 195183,
29	923	26.6	271	16	US-10-767-701-43225	Sequence 43225, A
30	921	26.5	224	16	US-10-767-701-32265	Sequence 32265, A
31	905	26.0	323	16	US-10-767-701-40999	Sequence 40999, A
32	887	25.5	303	16	US-10-425-115-188282	Sequence 188282,
33	833.5	24.0	522	16	US-10-437-963-109890	Sequence 109890,
34	794	22.8	156	15	US-10-424-599-175150	Sequence 175150,
35	784.5	22.6	187	15	US-10-424-599-175149	Sequence 175149,
36	737.5	21.2	279	15	US-10-424-599-161259	Sequence 161259,
37	730.5	21.0	280	15	US-10-424-599-278772	Sequence 278772,
38	715	20.6	262	16	US-10-425-115-260430	Sequence 260430,
39	697	20.1	265	15	US-10-424-599-246655	Sequence 246655,
40	674	19.4	227	16	US-10-425-115-366867	Sequence 366867,
41	668	19.2	714	9	US-09-749-589-4	Sequence 4, Appl
42	668	19.2	714	16	US-10-684-532-4	Sequence 4, Appl
43	662	19.1	751	9	US-09-795-693-14	Sequence 14, Appl
44	662	19.1	751	14	US-10-156-239-14	Sequence 14, Appl
45	662	19.1	751	14	US-10-199-485-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-10-437-963-140667
; Sequence 140667, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 140667
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41845C.1.pap
US-10-437-963-140667

Query Match 67.1%; Score 2330.5; DB 16; Length 657;

Best Local Similarity 70.6%; Pred. No. 8.4e-193;

Matches 441; Conservative 83; Mismatches 100; Indels 1; Gaps 1;

QY 43 QVEVPPQPFKSKYSLKSTETFPDPLRQF-KNPKASKKFMGLQFFPIFEWAPKYTF 101

18 RVPMPAAKPFLETGNNKMETLPDDPFVVRRCGRRRAAALRYVFPFMEWAPSYTL 77

QY 102 QFLKADLIAGITIASIAIPOGISYAKLANLPDILGLYSFISDPLIYAMGSSRDLAGTV 161
Db 78 GTLKSDLIAGITIASIAIPOGISYAKLANLPVGLIYSFVPLVYAMGSSRDLAGTV 137
QY 162 AVCSILMGSMNAVDNEDPKLYLHLAFTATLFAGVFOAALGLPRLGLIVDFLSHATII 221
Db 138 AVASLIDIGSMLESEAAADPDALYLHVALTATFFAGVFOALLGVRLGLIVDFLSHATIV 197
QY 222 GFMGGAATVVCILQQLKSIILGLEHFTHGADIIISVMRSVFTQTHEWRWESAVLGCVFIFLL 281
Db 198 GFMGGAATVVCILQQLKGMFLGDHFTTATDLVSMSSVFSQTHLWRWESVVMGGFLFLL 257
QY 282 STRYFSKKRPRFFWVSAMAPLTSVLGSLLVYTHAEKHGVEVIGELKKGLNPPSLNLV 341
Db 258 ITRFFSKRPRFFWVSAAAAPLASVIITGSLLVYLTHAENHGIQVIGYLLKKGLNPPSATSIN 317
QY 342 FVSPYMTTAVKTIIVGIIISLAEGIAVGRSFAMYNKYNIDGNKEMIAIGTMNVVGSFTSC 401
Db 318 FSSPYMMLAKTGIIITGVIALAEGIAVGRSFAMFNKYNIDGNKEMIAIGTMNIVGSLTSC 377
QY 402 YLTGPFSSAVNNAAGCKTAASNIIMSLAVMLTLLFLTPLFHYTPLVVLSAIIVSAMLG 461
Db 378 YLTGPFSSAVNNAAGCKTAMSNVIMSVAVMTLLFLTPLFHYTPLVVLSAIIMSLAG 437
QY 462 LIDYEAAILFKVDKDFVVCMSAYIGVFGSVEIGLVIATIVISVLRVLLFTARPRTFVL 521
Db 438 LIDYPAAVHLWQVDKDFCVCLGAYLGVVFGSVEIGLVAVAGISILRVLLFVARPRTTVL 497
QY 522 GNIPNSVIYRNVEHYQNAKHVPCMLLEIDAPYFANASYLRERITRWIDEEERIKATG 581
Db 498 GNIPNTMIYRMDQYTAARQVGVLRVDSPIYFTNAYLRERITRWIDDEEQCKEG 557
QY 582 ETSLOQVIIDMGAVIDGTSISMLBEVKKITERRELQVLVLPVSEVMKLNKSKFQNH 641
Db 558 EMGIQIVLDMGAVGSDITSGTSMDELAKTLDRGLQIVLANPGSEIMKMLDSSKVLRA 617
QY 642 LGKKWIIYLTVEEAVGACNFNLRSK 666
Db 618 IGHIEWIFPTVGEAABCDVFMHSQK 642

RESULT 2
US-10-424-599-273958
; Sequence 273958, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273958
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89404C.1.pap
US-10-424-599-273958

Query Match 58.0%; Score 2015; DB 15; Length 414;
Best Local Similarity 94.5%; Pred. No. 9.8e-166;
Matches 397; Conservative 8; Mismatches 9; Indels 6; Gaps 1;

QY 261 QTHEWRWESAVLGCVFIFLLSTRYFSKKRPRFFWVSAMAPLTSVLGSLLVYTHAEKH 320
Db 1 QTHEWRWESAVLGCVCIFLLVLTRYFSKKRQPKFFWVSAMAPLTSVLGSLLVYTHAEKH 60

QY 321 GVEVIGELKGLNPESLTNLVFPVSPYMTTAVKTIIVGIIISLAEGIAVGRSFAMYNKYN 380
Db 61 GVQ-----KGLNPESLTNLVFPVSPYMTTAVKTIIVGIIISLAEGIAVGRSFAMYNKYN 114
QY 381 DGNKEMIAIGTMNVVGSFTSCYLTGTPFSSAVNNAAGCKTAASNIIMSLAVMLTLLFLT 440
Db 115 DGNKEMIAIGTMNVVGSFTSCYLTGTPFSSAVNNAAGCKTAASNIIMSLAVMLTLLFLT 174
QY 441 PLFHYTPLVLSAIIVSAMLGILIDYEAAILFKVDKDFVVCMSAYIGVFGSVEIGLVI 500
Db 175 PLFHYTPLVLSAIIVSAMLGILIDYEAAILFKVDKDFVVCMSAYIGVFGSVEIGLVI 234
QY 501 AIVISLVLAVLFIAPRPTFVLGNIPNSVIYRNVEHYQNAKHVPCMLLEIDAPIFANAS 560
Db 235 AIAISVLRLVLLFIAPRPTFVLGNIPNSVIYRNVEHYNAKHVPCMLLEIDAPIFANAS 294
QY 561 YLRERITRWIDEEERIKATGETSIQYVVIDMSAVGNIDTSGISMLBEVKKITERRELQL 620
Db 295 YLRERITRWIDEEERIKATGETSIQYVIDMSAVANIDTSGISMLBEVKKITERRELQL 354
QY 621 VLVPVSEVMKLNKSKFQNHLGKKWIIYLTVEEAVGACNFNLRSKTHPKKDETEGWNV 680
Db 355 VLVPVSEVMKLNKSKFQNHLGKKWIIYLTVEEAVGACNFNLRSKTHPKKDETEGWNV 414

RESULT 3
US-10-425-115-235264
; Sequence 235264, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 235264
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_146143C.1.pap
US-10-425-115-235264

Query Match 53.3%; Score 1853.5; DB 16; Length 695;
Best Local Similarity 56.3%; Pred. No. 2e-151;
Matches 346; Conservative 115; Mismatches 153; Indels 1; Gaps 1;

QY 42 HOVEYPPPPQPFPSKLSYLSKETFFPDPLRQPKNPKASKKMLGLQFFPIFEWAPKYTF 101
Db 70 YKVGVPPEKNLLAEISDAVKETFFADDPRLRQYKQDQPSKKIWLGLQHFPPVLEWSRHSYL 129
QY 102 QFLKADLIAGITIASIAIPOGISYAKLANLPDILGLYSFISDPLIYAMGSSRDLAGTV 161
Db 130 GKFKGDFIAGLTIASLICIPQDIDIGYSKLANLPAEVLGYSFVPLIYAMGSSRDIAIGV 189
QY 162 AVCSILMGSMNAVDNEDPKLYLHLAFTATLFAGVFOAALGLPRLGLIVDFLSHATII 221
Db 190 AVVSLLLGLTLQNEIDPKTHPLEYRRLAFTATFFAGVTOAALGFFRLGLIIFLFLSHAIV 249
QY 222 GFMGGAATVVCILQQLKSIILGLEHFTHGADIIISVMRSVFTQTHEWRWESAVLGCVFIFLL 280
Db 250 GFMGGAATVVCILQQLKGMFLGFIAGFTKSDIVSMKSVWGNVHHGWNWQITILGATFLAFL 309
QY 281 LSTRYFSKKRPRFFWVSAMAPLTSVLGSLLVYTHAEKHGVEVIGELKKGLNPPSLNLV 340
Db 310 LVAKYIGKRNKKLFWVSAIAPLTSVIITFFYITRADKHGVAIVKNIIRKGINPPSASLI 369
QY 341 VFPVSPYMTTAVKTIIVGIIISLAEGIAVGRSFAMYNKYNIDGNKEMIAIGTMNVVGSFTS 400

Db 370 YFTGPGYLATGFKIGIVAGMIGLTAIGRTFAALKDXYRIDGNKEMMALGTNNVIGSLTS 429
Qy 401 CYLTGTPSPRSVANNAGCKTAASNIIMSLAVMLTLLFLTPPLFHYTPLVLSAIIYSAML 460
Db 430 CYVATGSFSRSVANNAGCKTAVSNVMSIVVMMLTLLLIITPLPKYTPNALISIIISAVL 489
Qy 461 GLIDYEAAILHLPKVDKDFVFCVMSAYIGVGVFSVEIGLVIAIVISLRLVLLFIARPRTFV 520
Db 490 GLIDYESAYLIWKVDKLDFLACMGAPGVIFSSVEYGLLIAVAISLAKIILQVTRPRTVL 549
Qy 521 LGNIPNSVYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRERITRWIDEERIKAT 580
Db 550 LGNLPRTIYRNVEQYDATKVPGLVIRVDSAIYFTNSYVKERILRWLRMLRDEEQOQD 609
Qy 581 GETSLQYVIDMSAVGNIDTSGISMLLEVKKITERRELQVLVNPVSEVMKLNKSKFQN 640
Db 610 KLTKTFLVLDLSPVIDIDTSGIHALEELAKALEKKRIQLVLNPGPAVIQKRSAKFTD 669
Qy 641 HLGKKWIYLTVEBAV 655
Db 670 MIGEDNIFLTVDGAV 684

RESULT 4

US-10-424-599-207945
; Sequence 207945, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 207945
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(659)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29800C.1.pap
US-10-424-599-207945

Query Match 53.2%; Score 1850; DB 15; Length 659;
Best Local Similarity 55.9%; Pred. No. 3.8e-151;
Matches 359; Conservative 122; Mismatches 153; Indels 8; Gaps 4;

Qy 41 VHOVEPPPPQPFKSLKYSKETFPPDDPLRQFNKNPKASKKFMGLQFPFPIPEWAPKYT 100
Db 20 VHOVPPPHKSTQKLGRLKETFPFDDPLRQFNKPKASKKFMGLQFPFPIPEWAPKYN 79
Qy 101 FQPKADLIAGITIASIAIPQGISYAKLANLPILGLIYSSFIPPLIYAMGSSRDLAGVT 160
Db 80 LKLFKSDLVSLGTIASLIDPQMSYAKLASLPILGLIYSSFVPLVYAVLGSSKDLAVGP 139
Qy 161 VAGSLLMGMLNADVDPNEDPKLYHLAFTATLPAQVFOAALGLFRLGLIVDFLSHAT 220
Db 140 VSIASLVMSMLHQEVSPTTDPILFLQLAFTSTLFAGLFOALLGILRLGFIIDFLSKAIL 199
Qy 221 IGFMGGAATVVCLOOLKSIILGLEHFTHGADIIISVMSVFTQTHHEWESAVLGCVFIFL 280
Db 200 IGFMGGAATVVCLOOLKSLIGITHFTNQMGILIPMTSVFNHESWQTIIMGICFLVLL 259
Qy 281 LSTRYSKKRPRPFVWSAMAPLTSVILGSLVYFTTHAEKHGVSIVIGELKKGLNPPSLTNL 340

Db 260 LLARHVSIKPKLFWVSAGAPLWCVIISTLLVFAIKAOHNGISAIGKLOOGINPPSNMML 319
Qy 341 VFPSPMTTAVKTGIIVGIIISL-AEGIAVGRSFAMFKYNNIDGNKEMIAIGTMVNVGSFT 399
Db 320 LPHGSHLGVLMTGTLTGILSLTXEGIAVGRFAALKKNYKVDGNKEMMAIGFMNVGSFT 379
Qy 400 SCYLTGTPSPRSVANNAGCKTAASNIIMSLAVMLTLLFLTPPLFHYTPLVLSAIIYSAM 459
Db 380 SCYVTTGAPSRSAVNNAGAKTAVSNVMSVTVMTLLFLMPLFQYTPNVVLGAIVTAV 439
Qy 460 LGIDYEAAILHLPKVDKDFVFCVMSAYIGVGVFSVEIGLVIAIVISLRLVLLFIARPRTF 519
Db 440 LGIDLPAACNLTDKDFVVMVMTAFGLVLFISVOGGLALAVGLTKLILQITRPKT 499
Qy 520 VLGNIPNSVYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRERITRWIDEERIKAT 579
Db 500 MLGKTPCTDIYNLQYKEAVRIPGFLIILSIEAPINFAITYNLERTLWIEEEDNIKE 559
Qy 580 TGETSQYVIDMSAVGNIDTSGISMLLEVKKITERRELQVLVNPVSEVMKLNKSKFQ 639
Db 560 --QLSLRFLVLEMSAVSAVDTSGLSFELKATLEKKGVELVLNPLASVIEKLIKADEA 617
Qy 640 N-HLGKKWIYLTVEBAVGNACNENLRASKTNPKKDETEGWNV 680
Db 618 NDFIRADNLFTVGEAVA----SLSSAMKQSGSTITEGAHTI 655

RESULT 5

US-10-437-963-198336
; Sequence 198336, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198336
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94005C.1.pap
US-10-437-963-198336

Query Match 52.1%; Score 1811.5; DB 16; Length 666;
Best Local Similarity 53.8%; Pred. No. 8.3e-148;
Matches 340; Conservative 122; Mismatches 167; Indels 3; Gaps 3;

Qy 43 QVEVPPPPQPFKSLKYSKETFPPDDPLRQFNKNPKASKKFMGLQFPFPIPEWAPKYTQ 102
Db 36 EYVLSGRPRFPAEKLSDLAETFPDDPPFRGFGALPPARAWCAVKYFVPALDWVRYGLD 95
Qy 103 FLKADLIAGITIASIAIPQGISYAKLANLPILGLIYSSFIPPLIYAMGSSRDLAGVTA 162
Db 96 KFPDLAGITIASIAIPQGISYAKLANLPILGLIYSSFVPLVYAVLGSSNNLAVGTVA 155
Qy 163 VGSLLMGMLNADVDPNEDPKLYHLAFTATLPAQVFOAALGLFRLGLIVDFLSHAT 222
Db 156 AASLLASIIIEVAADENPQLQLFYTAAPFTGLFQALGVRLGLIVDFLSRSTIG 215
Qy 223 FNGGAATVVCLOOLKSIILGLEHFTHGADIIISVMSVFTQTHHEWESAVLGCVFIFL 282
Db 216 FNGGTAMIILOQFKLLGMKHTTKTDIISVLHSTHYRHEWKMOSAVLGLICFLFLMS 275


```
QY 378 YNIDGNKEMIAIGTMNVGSGFTSCYLTGTPGFRSAVNNAGCKTAASNIIMSLAVMLTLL 437
DB 127 YHIDGNKEMIAIGTMNIFGFTSCYLTGTPGFRSAVNNAGCKTAASNIIMSLAVMLTLL 186
QY 438 EUTPLPHYTPPLVLSAIIIVSAMGLIDYEAAILHLFKVDKDFDVCMGSAVGVVFGSVETG 497
DB 187 EUTPLPHYTPPLVLSAIIIVSAMGLIDYEAAILHLFKVDKDFDVCMGSAVGVVFGSVETG 246
QY 498 LVIAIVISLVRLLFTIARPTFVLGNIPNSVIYRNVEHYQNAKHVPGMLILBIDAPIYFA 557
DB 247 LVIAIAISLVRLLFIARP-DXVLGNIPNSVIYRNVEHYPNAXHVPGLMLILBIDAPIYFA 305
QY 558 NASYLREIRTRWIDEBEERIKATGETSLQYVIDMSAVGNIDTSGISMLBEVKKITERRE 617
DB 306 NASYLREIRKRWIDEBEERIKA-----GMFDMSAXCMLGEI----- 341
QY 618 LQLVLVNPVSEVMKLNKSKFQNHGKWIYLTVEEAVGACPNLRASKTNPKKDETCGW 677
DB 342 -----PRSEVMKLNKSKFLDELQGWYLTVEEAVGACPNMLHSHYKPNPMKDESGW 394
QY 678 NN 679
DB 395 NN 396

RESULT 10
US-10-437-963-179495
; Sequence 179495, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179495
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(653)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76952C.1.pep
US-10-437-963-179495
```

Query Match 49.6%; Score 1723.5; DB 16; Length 653;
Best Local Similarity 53.2%; Pred. No. 3.4e-140;
Matches 328; Conservative 119; Mismatches 165; Indels 5; Gaps 2;

```
QY 42 HQVEVPPPPFFKSLKYSLKETFPPDDPLRQFNKPKASKKFMGLQFPFPIFEWAPKYTF 101
DB 28 YNVGAPPKNLLAEFAGTGVKETFSDPMRRYKQDQPSRKLMLALQHFVFEWGRQYTL 87
QY 102 QFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSSFIPLLIYAMGSSRDIAVGTV 161
DB 88 AKFPGDLIAGLTIASLIVIPDQIGYAKLANLPPIGLHSSFVPPLIYALMGTSRELAWGPV 147
QY 162 AVGSLMGMSLNAVDNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLSHATII 221
DB 148 AVISLLGLTLQREIDSXKNPLDYRLATPATTFAGVTQAALGFCRLGFIIFLSHAII 207
```

```
QY 222 GFMGAATVVCLOQLKSLGLEHFTHGADIIISVMESVFTQTH---EWRWESAVLGCVFIF 278
DB 208 GFMGAATIIAQQQLKGLFIANFTKTDIIISVMKSVGNVHHGQXQMNQIILIGASFLA 267
QY 279 FLLSTRYFSKKRPRFFWVSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGLNPPSLIT 338
DB 268 FLLVA--XGKKKKLFWVPALPLISVIISTLFVYITRADKQGVAVIKNVKKGINPPSAS 325
QY 339 NLVFSVPMTTAVKTVGVIGIISLAEGIAVGRSPAMYKNYINIDGNKEMIAIGTMNVGSGF 398
DB 326 LIFFTPYLLKGFKIGVAGMISLTEAIAVGRTEAGLNDYQIDGNKEMIALGTMTNVVSGM 385
QY 399 TSCYLTTPGFRSAVNNAGCKTAASNIIMSLAVMLTLLFTPLPHYTPPLVLSAIIYSA 458
DB 386 TSCYIATGCFARSAVNCMAGGKTPMSNIIVMSVTLALLWITPLPKYTPNATISSIIISA 445
QY 459 MLGLIDYEAAILHLFKVDKDFDVCMGSAVGVVFGSVETGLVIAIVISLVRLLFIARPT 518
DB 446 VLGLDFPESAVLIWKVDKLDPMACLGAFGLVIFSSVEYGLLIJAVVISLIKVLHVTRPT 505
QY 519 FVLGNIPNSVIYRNVEHYQNAKHVPGMLILBIDAPIYFANASYLRERTRWIDEBEERIK 578
DB 506 ALLGNLPRTIYRNVEQYPEATKVPGLIIVRDSAIYFTNSNYVKERMLRLRDEEHQK 565
QY 579 ATGETSLQYVIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKLNKSKF 638
DB 566 EQLPKIBFLIVDLSPVNDIDTSGIHAFKELLRTLEKQIQILIFANPGAAVIQKURSAKF 625
QY 639 QNHGKWIYLTVEEAV 655
DB 626 TELIGEKKICTVGDAV 642
```

```
RESULT 11
US-10-425-115-366870
; Sequence 366870, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 366870
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(662)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97750C.1.pep
US-10-425-115-366870
```

Query Match 49.6%; Score 1722.5; DB 16; Length 662;
Best Local Similarity 53.9%; Pred. No. 4.3e-140;
Matches 343; Conservative 110; Mismatches 168; Indels 15; Gaps 6;

```
QY 41 VHQVEVPPPPPPFFKSLKYSLKETFPPDDPLRQFNKPKASKKFMGLQFPFPIFEWAPKYT 100
DB 35 VHKVAPPAPASPAASKMKVRVKETFPFDDPFRAFQKQPQCTQMLMAVRYLFPILDWVPSYS 94
QY 101 FOFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSSFIPLLIYAMGSSRDIAVCT 160
DB 95 LSLFKSDLVAGITIASLAIPQGISYAKLANLPILGLYSSFIPLLIYAMGSSRDIAVGP 154
QY 161 VAVGSLMGMSLNA-VDPNED---PKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFL 215
```

Db 155 VTSSLMGKPCXACRPHRGADAVPAAGLH-----XTLFAGLVQAASLGLILGFFVIDFL 210
Qy 216 SHATIGFGWGAATVVCLOOLKSLGLEHFTTGADIIISVMRSVFTOTHEWRWESAVLGCY 275
Db 211 SKATLVGFWAGAAIIVALQOLKALJGIVFTTGMGIVPVNASVFHHTSWSQTILMGVC 270
Qy 276 FIFLLSTRYFSKRRPRFFWWSAMAPLTSVILGSLVLYFTHAEKKGVEVIGELKKGKLNPP 335
Db 271 FLVFLLSARHVSIRWPKLFWWSACAPLASVTISTLLVFLPKAQNHGISIIGQKCGLNRP 330
Qy 336 SLTNLVFVSPYMTTAVKTGIVGIIISLARGIAGVRSFAMKYKNYIDGNKEMIAIGTMNVV 395
Db 331 SWDKLLFDYAILGLTWKTLGLTGTIIISLREGIAGVGTFFASLKQYQIDGNKEMIAIGLMNVV 390
Qy 396 GSFTSCYLTGTPRSANVYAGCKTAASNIIMSLAVMLTLLFLTPHYTLPVLVLSAII 455
Db 391 GSTCSYVTGTFAPRSANVHAGCKTAMNSVIMALTVMVTLFLMPLFYTTNVVLGAIL 450
Qy 456 VSAMGLIDYEAAIHLFKVDKDFVVCMSAYIGVWFGSVVEIGIIVIAVISLVRLLFIAR 515
Db 451 IAAVIGLIDFPAYVHLWKMDKDFLCVCAPAGVIFISVQEGIAIAGVISIFRVLWQITR 510
Qy 516 PRTFVLGNTPNSVIYRNVEHYQNAKHVPKMLIETDAPYIFANASYLRERITRWIDEBE 575
Db 511 PKMWQGNIKGTDIYRDLHHYKQAVSGVFLIATIEAPINFANSNYLNERIKRWI--EE 568
Qy 576 RIKATGETSLOVVIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKLNK 635
Db 569 SFEQDKHTELHFLIIDLASVAPALDTSGLIAPLIDIKKSIKRGLELVLPVNPTEGXHGNTT 628
Qy 636 SK--FQNLHGKWKIYLTVEEAVGACNFNLRSKNTNP 669
Db 629 CRGXXTILGQACIYTTGEAI--ASLSALAKWTKP 662

RESULT 12

US-10-437-963-174545
; Sequence 174545, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174545
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72475C.1.pap
US-10-437-963-174545

Query Match 45.0%; Score 1563.5; DB 16; Length 659;
Best Local Similarity 47.7%; Pred. No. 2.6e-126;
Matches 299; Conservative 121; Mismatches 194; Indels 13; Gaps 3;
Qy 50 QPEFKSLKYSKTEFFDDPLRQFNKPKASKKFMGLQFFFPFIFEWAPKYTFQFLKADLI 109
Db 29 RPEGAARALKDTLFPDDPFRGLGMPARRAWRVARYFPALDWAGYSAASFWDLL 88
Qy 110 AGTTIASLAPQGISYAKLANLPPIILGLYSSFPPIIYAMGSSRDLAVGTAVGSLMG 169

Db 89 AGVTIASLIPQGISYATLAGIPPVIGLYSCFVPLVAVMGSSRNGLGVGPVATSLVA 148
Qy 170 SMLSNVADPNEDPKLYLHLAFATATLFAGVFQAALGLFRGLGLVDFLSHATIIIFMGGAAT 229
Db 149 SVIGCKVRASDDORLYTQLVFTSAFTGTVLQAALGLRLGILVDPMRSRPAITGFMGTAI 208
Qy 230 VVLCLOOLKSLGLEHFTTGADIIISVMRSVFTOTHEWRWESAVLGCYFFFLSTRYFSKK 289
Db 209 VTMLOOLKGLFGMTHTTKTDIVSVLRYIFHNTHOMQSTVLGVCFILFLVFTQVRRR 268
Qy 290 RPRFFWWSAMAPLTSVILGSLVLYFTHAEKKGVEVIGELKKGKLNPPSLTNLVFVSPYMTT 349
Db 269 RPKLFWWSAMSPLLVVVVCVFSFLIKGHKGIPVIGLKRGINFSSISQLQFQPEYGV 328
Qy 350 AVKTGIVVGIISLARGIAGVRSFAMKYKNYIDGNKEMIAIGTMNVVSGTSCYLTGTPFS 409
Db 329 AMKAGFVSGMLAALASGVAVGRSFAAMKKERIDGNKEMVAFGLMNLIGSTSCYITTGAFS 388
Qy 410 RSVNNAAGCKTAASNIIMSLAVMLTLLFLTPHYTLPVLVLSAIIIVSAMGLIDYEAAI 469
Db 389 KTVNHYHAGCRTAMSNVCMALVVALAPLFRHTPLVALAAIITSSMLGLVKHREIR 448
Qy 470 HLFKVDKDFVVCMSAYIGVWFGSVVEIGIIVIAVISLVRLLFIARPTFVLGNTI----- 524
Db 449 RLYEVDKADFVCAALIGVVFSTMITGLGVAVALSVLRALUHVAPSTSKUGRVSCGS 508
Qy 525 ----PNSVIYRNVEHYQNAKHVPKMLIETDAPYIFANASYLRERITRWIDEBEERIK 579
Db 509 AGAADDHAFCDVQPGNATAPSLVLQVAGSPVCFANAEYLRERIAWVEDEB---KA 565
Qy 580 TGETSLOVVIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKLNKSKFQ 639
Db 566 VAGEDLLVVLIDIGVTAIDSPGIEMLRVHGLSRKGMKMAVTNPRMAVASKVLVSLGLA 625
Qy 640 NHLGKWKIYLTVEEAVGACNFNLRSK 666
Db 626 ELVGESWMFLSNGDALAACRYTLQSGK 652

RESULT 13

US-10-424-599-195017
; Sequence 195017, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 195017
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_18127C.1.pap
US-10-424-599-195017

Query Match 43.2%; Score 1500; DB 15; Length 653;
Best Local Similarity 48.0%; Pred. No. 8.1e-121;
Matches 297; Conservative 126; Mismatches 178; Indels 18; Gaps 7;
Qy 48 PPQEFKSLKYSKTEFFDDPLRQFNK-----PASKKFMGLQFFFPFIFEWAPKYT 100
Db 38 PRSP-WRVVADSVSKT-----ISHYKHLSSLIQPCTCTLLSLVLQVVPFILA WGRNYT 90
Qy 101 FOFLKADLIAGTITIASLAPQGISYAKLANLPPIILGLYSSFPPIIYAMGSSRDLAVGT 160
Db 91 ATKFRKDLIAGTITIASLCPISQIGYATLAHLDPQYGLYTSVVPPLIYAVMGTSRITAIGP 150

Qy	455	IVSAMLGLIDYEAAILHFKVDKDFVVCMSAYIGVWFGSVBEIGLVIATIVISVLRVLLFIA	514
Db	179	IVSAMLGLVDFGAALHLWRVDKDFVCAGAYLGVWFGSVBEIGLVAVAVSLLRVLRFGA	238
Qy	515	RPRTFVLGNIPNSVIYRNVEHYONAKHVPQMLILEIDAPIYFANASYLRERITRWIDEE	574
Db	239	RPRTTVLGNIPGTWVYRRMDQYAAATVPGVLVLRVDAPVYFANASYLRERISRWIDEE	298
Qy	575	ERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELOLVLPVSEVMKKLN	634
Db	299	ERTKSQEMGVRVVLDMGAIGSIDTSGTSMLELKNKSLDRRGMQIVLANPGSEIMKKLD	358
Qy	635	KSKFQNLGKKIYLTVEEAVGACNFENLRASKTNPKKD	672
Db	359	SSKVLQIGHEWVFPPTVGEAVASCDYVLHSHKPGMAKD	396

Search completed: September 12, 2005, 19:14:19
Job time : 1410 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 12, 2005, 18:42:11 ; Search time 251 Seconds
(without alignments)
4432.941 Million cell updates/sec

Title: US-10-762-049-18
Perfect score: 3475
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDDEGWNV 680

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DBV=xlp
-Q=/cgn2_1/USPTO.spool_p/us10762049/runat_12092005_120740_7209/app_query.fasta_1.839
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10762049@cgn_1_1_105@runat_12092005_120740_7209 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3475	100.0	2449	US-09-720-317A-17	Sequence 17, Appl
2	2164	62.3	1981	US-09-720-317A-3	Sequence 3, Appl
3	1901	54.7	2067	US-09-720-317A-15	Sequence 15, Appl
4	1844.5	53.1	2279	US-09-720-317A-1	Sequence 1, Appl
5	1799	51.8	2311	US-09-720-317A-19	Sequence 19, Appl
6	963.5	27.7	2022	US-09-720-317A-21	Sequence 21, Appl
7	731.5	21.1	780	US-09-720-317A-7	Sequence 7, Appl
8	687.5	19.8	1791	US-09-252-991A-7241	Sequence 7241, Ap
9	676.5	19.5	1767	US-09-252-991A-7184	Sequence 7184, Ap
10	675.5	19.4	1240	US-09-720-317A-5	Sequence 5, Appl
11	651	18.7	2441	US-09-785-381-4	Sequence 4, Appl
12	650	18.7	4113	US-09-785-381-2	Sequence 2, Appl

13	643.5	18.5	1773	4	US-09-902-540-7746	Sequence 7746, Ap
14	643.5	18.5	4854	4	US-09-902-540-768	Sequence 768, Ap
15	641.5	18.5	2832	4	US-09-949-016-4516	Sequence 4516, Ap
C 16	625	18.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 17	619	17.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
18	612	17.6	1818	4	US-09-602-787A-615	Sequence 615, App
19	586.5	16.9	2858	4	US-09-949-016-5349	Sequence 5349, Ap
20	586.5	16.9	2882	1	US-08-424-567-1	Sequence 1, Appli
21	586.5	16.9	2882	2	US-08-711-928-1	Sequence 1, Appli
22	586.5	16.9	2882	3	US-09-184-937-1	Sequence 1, Appli
23	578	16.6	2682	4	US-09-614-221A-222	Sequence 222, App
24	577	16.6	4327	4	US-09-949-016-5827	Sequence 5827, Ap
C 25	571.5	16.4	31147	4	US-09-596-002-25	Sequence 25, Appl
26	567	16.3	2487	4	US-09-248-796A-6643	Sequence 6643, Ap
27	548.5	15.8	1992	4	US-09-875-811-5	Sequence 5, Appli
28	548	15.8	1971	4	US-09-875-811-9	Sequence 9, Appli
29	548	15.8	2420	4	US-09-875-811-13	Sequence 13, Appl
30	545.5	15.7	2040	4	US-09-875-811-1	Sequence 1, Appli
31	489	14.1	1851	4	US-09-543-681A-249	Sequence 249, App
32	485.5	14.0	1818	4	US-09-875-811-7	Sequence 7, Appli
33	485	14.0	1797	4	US-09-875-811-11	Sequence 11, Appl
34	482.5	13.9	1866	4	US-09-875-811-3	Sequence 3, Appli
35	464	13.4	8774	4	US-09-949-016-16258	Sequence 16258, A
36	451	13.0	2513	4	US-09-795-927-6	Sequence 6, Appli
37	449	12.9	1608	4	US-09-795-927-8	Sequence 8, Appli
38	449	12.9	1911	4	US-09-252-991A-5234	Sequence 5234, Ap
39	449	12.9	1911	4	US-09-252-991A-5199	Sequence 5199, Ap
40	426.5	12.3	1419	4	US-09-902-540-6127	Sequence 6127, Ap
41	426.5	12.3	2070	4	US-09-902-540-286	Sequence 286, App
C 42	402.5	11.6	1230230	4	US-09-438-185A-1	Sequence 1, Appli
C 43	389.5	11.2	1230025	4	US-09-198-452A-1	Sequence 1, Appli
C 44	372	10.7	1116	4	US-09-252-991A-7460	Sequence 7460, Ap
45	369	10.6	1701	4	US-09-543-681A-372	Sequence 372, App

ALIGNMENTS

RESULT 1
US-09-720-317A-17
; Sequence 17, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 2449
; TYPE: DNA
; ORGANISM: Glycine max
US-09-720-317A-17

Alignment Scores:			
Pred. No.:	0	Length:	2449
Score:	3475.00	Matches:	680
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-762-049-18 (1-680) x US-09-720-317A-17 (1-2449)

QY	1	HisGluLeuAlaArgThrLeuSerTyrIleThrHisIleCysLeuLeuArganThrIle	20
DB	2	CACGAGCTAGCTGGCACATTAAGTTATATACACATATTGCTTGCTTAATACTATT	61

QY 21 IleGluAspMetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArg 40
Db 62 ATTGAAGATATGGGAGGTGATGATTAATGATGATCCCTTTGGCATGAACAACCTTTGAGAGA 121
QY 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
Db 122 GTGCACCAAGTCAGAGTTCACCGCCACACCGCCGCTTTTCAAGTCTCTAAGTACTCTTTG 181
QY 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db 182 AAGGAGACTTCTTCCTCGATGACCTTTGAGGACGTTCAAGAACCAAGCAGCTTCCAAG 241
QY 81 LysPheMetLeuGlyLeuGlnPhePheProIlePheGluThrAlaProLysTyrThr 100
Db 242 AAGTTCAATGCTTGGCCTTCAGTCTCTTCTCCCATTTTGAATGGGCTCCCAATAACACC 301
QY 101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
Db 302 TTTTCAGTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCT 361
QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140
Db 362 CAGGGCATCAGTTATGCCAAGCTCGGCAACCTCCCTCCAATTCCTGGACTATATTCGAGC 421
QY 141 PheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db 422 TTTATACCACTTGAATGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 481
QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
Db 482 GTGCGGGTGGATCGCTCTGATGGGTTCGATGTTCGATTAATGCCGCTTCATCCCAATGAA 541
QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
Db 542 GACCCAAAGCTTTACCTCCACCTGGCTTTCACAGCTACATATTTCTGCTGGTGTGTTTTCAG 601
QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
Db 602 GCTGCTTGGGCTGTGTTAGTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 661
QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
Db 662 ATAGGTTTCATGGAGGAGCAGCAGCGTGGTGGTCTGCAGCAACTTAANAATCGATTCCT 721
QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
Db 722 GGCCTTGAGCATTTCCACCATGGAGCTGATATCATATCATGATGCGCTCTGTTTTCACC 781
QY 261 GlnThrHisGluThrArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
Db 782 CAAACTCATGAGTGGAGGTGGGAAGTGTGTGTAGGATGTCTTCATTTTCTTTCCTC 841
QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
Db 842 CTTAGCACAGATACTTCAGCAAAAACGACCCAGGTTTTTTGGGTGTGAGCAATGGCG 901
QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db 902 CCATTCAGCTCCCTTATATTTGGGAAGTCTCTTGGTTTATTTCACTCACGCCGAGAGAC 961
QY 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProSerLeuThrAsnLeu 340
Db 962 GGTGTTGAAGTATAGGAGAACTGAAGAAGGGTTTGAATCCACCATCTCAAAATCTG 1021
QY 341 ValPheValSerProTyrTrpThrAlaValLysThrGlyIleValValGlyIleIle 360
Db 1022 GTATTTGTGCGCTTACATGACTACAGCTGTCAAAACTGGCATTTCTGTTGGCATCAT 1081
QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
Db 1082 TCACCTTGGGAGGAATAGCAGTAGGAAGCTTTGCAATGTATAAAAATTACAATATT 1141
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400

Db 1142 GATGCCAACAAGAGATGATAGCTATTGGACCATTGAACGTAGTTGGTTCTTTCACTCT 1201
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
Db 1202 TGCTACCTCACAAACAGGACCACTTTTCGGGTTCGGCTGTGAACATAAAGCTGGATGCAAG 1261
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db 1262 ACAGCAGCTTCCAACATTAATAATGTCACCTTCAGTAATGTTGACATTTGTTATTCCTGACA 1321
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db 1322 CCCTTGTTCATTACACTCCCTCGTGGTGTCTATCAGCTATTATCGTATCTGCAATGCTT 1381
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 1382 GGACTCATAGATTATGAAGCAGCCATCCATCTATTAAAGGTTGACAAATTTGACTTTGTG 1441
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db 1442 GTGTCATGAGTGCATACATTGGCGTGTCTTTGGCAGTGTGAAAATTTGGCTTAGTCATA 1501
QY 501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520
Db 1502 GCTATTGTAAATCTGTACTTCGGGTACTTCTATTATTATGCAAGGCCAAGACATTCGTT 1561
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1562 TTTGGCAACATTTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAA 1621
QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db 1622 CATGTTCTGGAAATGCTAATTTCTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAGC 1681
QY 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db 1682 TATTTAAGAGAAAGGATCACAGGTGGATTGATGAAGAAGAAGAAGAAATTAAGCTACA 1741
QY 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db 1742 GGGGAGACTAGTTTGCAGTATGTTAATAATGATATGATGCTGTGTTGGAACATTGATACA 1801
QY 601 SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db 1802 AGTGAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTT 1861
QY 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db 1862 GTTTTGGTCAATCCTGTAAGTGAAGTGAAGAAACTGAAACAAATCGAAGTTCCANAAT 1921
QY 641 HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPhe 660
Db 1922 CATTTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGGCGCTTGGAGCATGCAACTTC 1981
QY 661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrAsnAsnVal 680
Db 1982 AATCTACGTGCAAGCAAAACGAAACCAAGAAAGATGAACACAGAGGGTTGGAAACAATGTG 2041

RESULT 2

US-09-720-317A-3
; Sequence 3, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998

```

; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-3

```

Alignment Scores:

Pred. No.:	2.17e-258	Length:	1991
Score:	2164.00	Matches:	403
Percent Similarity:	86.0%	Conservative:	86
Best Local Similarity:	70.95%	Mismatches:	79
Query Match:	62.27%	Indels:	0
DB:	4	Gaps:	0

US-10-762-049-18 (1-680) x US-09-720-317A-3 (1-1981)

QY	105	LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer	124
DB	8	GAGTCCGACCTCATCGCGGATCACCATCCAGCTCCGCAATCCCGAGGCGATCAGC	67
QY	125	TyrAlaLysLeuAlaAsnLeuProPheIleLeuGlyLeuTyrSerSerPheIleProPro	144
DB	68	TACGCCAGCTCGCCAAACTCGCGGCCGTGCTCGACTTCTACTCGAGCTTCGTGCCGCG	127
QY	145	LeuIleTyrAlaMetMetGlySerSerAspLeuAlaValGlyIleThrValAlaValGly	164
DB	128	CTGTGTATACGCCTCATGGGAGCTTCCAAGACACTGGCGGTGGGACGGTGGCGGTGGCG	187
QY	165	SerLeuLeuMetGlySerMetLeuSerAsnAlaValaAspProAsnGluAaspProLysLeu	184
DB	188	TCGTGCTCATCAGCTTCCATGCTGGCAGGAGGTGTCGCGCGCGGAACACCCCGTGCTC	247
QY	185	TyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuGly	204
DB	248	TACCTGCACCTCGCCTTACGCCACACTTCTTTCGCGCGGTCTTCCAGGCTCGCTCGGC	307
QY	205	LeuPheArgLeuGlyLeuIleValaAspPheLeuSerHisAlaThrIleIleGlyPheMet	224
DB	308	CTCCTCAGGTGGGGTTTCATCGTGACCTGCTGTCACGCGACATCGTGGGGTTTCATG	367
QY	225	GlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis	244
DB	368	GCGGCGCGCGCACGGTGTGTGCTGCAGCAGCTGAAGGCGATCTGGGCGCTGCTCCAC	427
QY	245	PheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlu	264
DB	428	TTTCAACCTCCACCGACGTCTCGTTCATGGAATCCGTCTTTCAGCCAGACACACCAG	487
QY	265	TrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeuLeuSerThrArg	284
DB	488	TGGCGTGGGAGAGGGTCTGCTGCTGGCTGCGGCTTCTCTTCTCTCTCTGTCACCGG	547
QY	285	TyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSer	304
DB	548	TTATCATCAGCAAGAGCGTCCCAAGCTGTTCTGGATCTCCGCGCGCGCGCTTGACGTCC	607
QY	305	ValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyValGluVal	324
DB	608	GTCTGCTCGGAGCGTTCTGTGTACCTCATCGCACGCTGAAACACACGCGCATCGAAGTG	667
QY	325	IleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSer	344
DB	668	ATCGGTTACTCGAAGAAAGGCTGAATCCACCGTGGTGACAGCGCTGCAATTTCTCACCG	727
QY	345	ProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGlu	364
DB	728	CCCTACATGATCTCTCGCGCTCAAGACTGGGATCATACCGGCGTTCATTCGCTCGCGAA	787
QY	365	GlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLys	384
DB	788	GGAAATCGCGTGGGAGAGAGCTTCGCAATGTTCAAGAACTACACATGACGGCAACAAG	847

/ Sequence 1, Application US/09720317A
 / Patent No. 6696292
 / GENERAL INFORMATION:
 / APPLICANT: Stephen M. Allen
 / APPLICANT: Saverio C. Falco
 / APPLICANT: Catherine J. Thorpe
 / TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
 / FILE REFERENCE: BB-1167
 / CURRENT APPLICATION NUMBER: US/09720,317A
 / CURRENT FILING DATE: 2000-12-21
 / PRIOR APPLICATION NUMBER: 60/092,833
 / PRIOR FILING DATE: 14-07-1998
 / NUMBER OF SEQ ID NOS: 31
 / SOFTWARE: Microsoft Office 97
 / SEQ ID NO 1
 / LENGTH: 2279
 / TYPE: DNA
 / ORGANISM: Zea mays
 / US-09-720-317A-1

Alignment Scores:

Pred. No.: 1 69e-218 Length: 2279
 Score: 1844.50 Matches: 355
 Percent Similarity: 74.25% Conservative: 115
 Best Local Similarity: 56.08% Mismatches: 158
 Query Match: 53.08% Indels: 5
 DB: 4 Gaps: 3

US-10-762-049-18 (1-680) x US-09-720-317A-1 (1-2279)

QY 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
 Db 189 GTGCACAAGGTGGCGCGCCACCGCGCGAGCAGCGAGCAAGATGAAGGTGAGGGTG 248
 QY 61 LysGluThrPhePheProLysAspProLeuArgGlnPheLysAsnLysProLys 80
 Db 249 AAGGAGACCTTCTCCCGACGACCGGCTTCCGCGGCTTCAAGGGGCGCGCGGAGC 308
 QY 81 LysPheMetLeuGlyLeuGlnPhePhePheProLysGluThrAlaProLysTyrThr 100
 Db 309 CAGTGGCTCATGGGTGAGTACCTCTTCCCATCTCGACGTGGGTGCGAGTACTCC 368
 QY 101 PheGlnPheLysAlaAspLeuAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
 Db 369 TTGTGCTCTTCAAGTCCGACCTCGTCCGCGGCTCACCATTGCCAGCTCGCCATCT 428
 QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProLysLeuGlyLeuTyrSer 140
 Db 429 CAGGGCATTAGCTACGCGAAGCTGGCAAGCTTCCCTCCCATATCGGGCTGTATTCGAGC 488
 QY 141 PheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
 Db 489 TTGTGCGCGGATGTGACGGGTGCTGGGAGCTCCCGTACCTGGCGGTGGCGCCG 548
 QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
 Db 549 GTGTGATCTGTGCTGATCATGGGTTCATGCTCGCGGCGCGCTGAGCCCACTGCG 608
 QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
 Db 609 GAGCCGACGCTGTTCTTCGAGCTGGCTTCCACCTCCACCTGTTCCGCGGGCTGGTGCAG 668
 QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
 Db 669 GCCTCCCTGGCATCTCAGGCTCGGCTTCGTCATCGACTTCTGTCCAAGGCGAGCTG 728
 QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
 Db 729 GTGGGTTCATGCGCGCGCGCCCATCATCTGTTGGCGCTGCAGCAACTCAAGGGGTGCTG 788
 QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
 Db 789 GGCATCGTCCACTTCAACCCGAGATGGGCGATCGTCCCGATCATGCGCTCCGCTCTTCCAC 848

QY 261 GlnThrHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
 Db 849 CACACGAGGAGTGGTGGCGAGCATCTCATGGGGTCTGCTCTCTCTCTCTCTCTCT 908
 QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTyrPheValSerAlaMetAla 300
 Db 909 CTGTGGCGGAGCATGTGAGCATCATGATGGCAAGCTTTTCTGGGTTTCTGGCGTGGCG 968
 QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
 Db 969 CCCCTGGCATCGGTCAACATCTCGACGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1028
 QY 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
 Db 1029 GGCATCAGCATCATTTGGGCGAGCTCAAGTCGGCGCTGAATCGCCCTCTCTGGGCAAGCTC 1088
 QY 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
 Db 1089 CTGTTTGACACGGCGTATTTAGGCTTCAACATGAAGTGGCTTGTCTCAACGAGTCAACCAT 1148
 QY 361 SerLeuAlaGluGlyIleAlaValAlaGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
 Db 1149 TCATGACGGAAGGATAGCGGTGTGTAGAACATTTGCTCTCACTCAAGGACTACCAAGATA 1208
 QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValIleGlySerPheThrSer 400
 Db 1209 GATGGAAACAGGAGATGATGCGCATAGGTTGATGAATGTTGTTGGTCTCTGCACATCA 1268
 QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
 Db 1269 TGCTACGTAAACAACAGGTGCTCTCCGCTCTGTGTAAACCAACACACGCGCTGCAAG 1328
 QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
 Db 1329 ACTGCCATGTCCAAAGTATCATGCGCTGATGTGATGCTGCTGCTGCTGCTGCTGCTGCT 1388
 QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
 Db 1389 CACCTGTGTGTGTACACACCCAACTGTTCTCGGAGCGATCATATATCCCGGGTGTATC 1448
 QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
 Db 1449 GGCCTGATGATTTCCCGCGGTGTACACATCTGGAAGATGGAAGATGATGATTTCTG 1508
 QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
 Db 1509 GTGTGCGTTTGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1568
 QY 501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520
 Db 1569 GCGGTGTGTATATCTATATTAGGTGTGTATGATGATGATGATGATGATGATGATGATGAT 1628
 QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
 Db 1629 CAAGGGAACATCAAGGGAGCTGATATTTACAGAGACCTGCATCACTACAAAGGAGGCCAA 1688
 QY 541 HisValProGlyMetLeuLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
 Db 1689 AGAGTTTCTGGGTCTTGTATCTTGGCCATTGAAGCAGCATTAACCTTCCCAACTCCAAC 1748
 QY 561 TyrLeuArgGluArgIleThrArgTyrPheAspGluGluGluArgIleLysAlaThr 580
 Db 1749 TACCTGAATGAAGATTAAGATGATA-----GAGGAAGATCTTTTGAACAGGAT 1802
 QY 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
 Db 1803 AAACATCTGAATCCATTTTCAATACTTGGATCTGTGAGTGTCTCTCTCTCTCTCTCTCT 1862
 QY 601 SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620
 Db 1863 AGTGGCATAGCGTCTCTCATTTGACATAAAGAAATCAATAGAGAAACGTTGGTCTGGAGCTT 1922


```

Db      1687 AAATTTACTCTCGGGTGGTGATTTGTGAGGGTTGATTCTGCTATTATTTTCCAACTCT 1746
Qy      560 SerTyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleValAla 579
Db      1747 AATTAGCTCGGAAGAAGAAATCTTAGGTGGTGCACAGACGAAGAAGACAGAGTAAAGCA 1806
Qy      580 ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp 599
Db      1807 GTGGGATTGGCTAAATCAGTTTCTCTGATTGTGGAATGTGCGCGGTCTACGACATCGAT 1866
Qy      600 ThrSerGlyIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGln 619
Db      1867 ACAAGCGGCATACATCTCTTGAAGATCTATACAAGAATCTTCAGAAAAAGATATGCAG 1926
Qy      620 LeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln 639
Db      1927 CTCATTCTGTGCAATCTGTGTTCCGTCGTCATAGAAAAAAGTGCAGCGTCCGAAGCTCAC 1986
Qy      640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db      1987 GAGCACATTGGAAGCAGCAATATATTCTCGCGGTCTCTGACGCTGTGCGATTCTGTACG 2046
Qy      660 PheAsnLeuArgAlaSerLysThrAsnPro-LysLysAspGluThrGluGlyTyr 677
Db      2047 -----ACGAAGTCGATGCGAGAACCGTGAGCGAAGTAGTTTCGGAGGAATGG 2092

RESULT 6
US-09-720-317A-21
; Sequence 21, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-720-317A-21

Alignment Scores:
Pred. No.: 6,81e-109 Length: 2022
Score: 963.50 Matches: 206
Percent Similarity: 56.30% Conservative: 129
Best Local Similarity: 34.62% Mismatches: 243
Query Match: 27.73% Indels: 17
DB: 4 Gaps: 7

US-10-762-049-18 (1-680) x US-09-720-317A-21 (1-2022)
Qy      91 ProIlePheGluTrpAlaProLysTyrThrPhe---GlnPheLeuLysAlaAspLeuIle 109
Db      66 CCGTGTTCGGTGGATGGGAGCTACAGATGAAGGAGGAGTTCACGCGCGACCTCGCC 125
Qy      110 AlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAla 129
Db      126 GCGCGCATCACTGTGCGGTGATGTGCTGCGCTCAGGCAATGTCATATCAAGCTGGCT 185
Qy      130 AsnLeuProIleLeuGlyLeuTyrSerPheIleProLeuIleTyrAlaMet 149
Db      186 GGGCTTCACCAATTTAGGCTCTACACAGGCTTTGTCCCATCTATTTCTACCGGATT 245
Qy      150 MetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuMetGly 169
Db      246 TTTGGTCTCTACGACAAATTAGCAGTAGTCCAGTGGCACTTGTCTCTCTGCTAGTGC 305

```

```

Qy      170 SerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAla 189
Db      306 AATGTTCTTGGGGTATAGTT-----AATTCACTAGTACGCTGTACACGAATTAGCC 359
Qy      190 PheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuLeuGlyLeuPheArgLeuGly 209
Db      360 ATATTATTGGCATTCATGTTTGGAAATCTGGAATGCTTGGCATTTGTAAGACTTGGC 419
Qy      210 LeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThr 229
Db      420 TGGCTTATTTCGTTTTCATTAGCCATCTGTAATATCTGGAATCTCCTACAGCTTCGGGCATC 479
Qy      230 ValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAla 249
Db      480 GTAATTGGTTTGTCCCAATCAAGTATTCTTGGGTACAGT---GTTACAAGAAGTAGC 536
Qy      250 AspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGluSer 269
Db      537 AAAATTATACCACTTATTGAGAGTATAAATGTCTGGAATAGATCAGTTCTCTCGGCTCCA 596
Qy      270 AlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSerLysLys 289
Db      597 TTTGTAATGGGATCAGCGTTTCTGTTATTCTTCTAATAATGAAGAAAGCTAGGGAACA 656
Qy      290 ArgProArgPhePheThrValSerAlaMetAlaProLeuThrSerValIleLeuGlySer 309
Db      657 AATAAAAAATACGTTTCTCGAGAGCTTCTGTCCTCACTAACAGCTGTGTTCTTGGAA 716
Qy      310 LeuLeuValTyrPheThrHisAlaGlyLysHisGlyValGluValIleGlyGluLeuLys 329
Db      717 TTTGTTGTGMAAATTTTTCGCTCCAACCT-----GCCATATCAGTGTGAGTGAATACCG 770
Qy      330 LysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerPro----- 345
Db      771 CAAGGCTT-----CCCACT-----TTCCTCATTCCTCGAGATTGTGA 809
Qy      346 TyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGly 365
Db      810 CATCTGATGTCCCTAATGCCAATCTATCACTGCTGGTGTGCTATTTTGGAGTCT 869
Qy      366 IleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGlu 385
Db      870 GTTGGGATTGCTAAAGCGTTAGTCGGAAGAATGTTATGAGTTGAGCTCAAAACAAGAG 929
Qy      386 MetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThr 405
Db      930 TTATTGTCCTTGGCTTATCAAAATATATGCGGTTTCATTCTCTCTGCATATCCTGCTACA 989
Qy      406 GlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsn 425
Db      990 GGCTCTCTTTCTAGGTCTGCTGTGATCATGAAAGCGGGGAAAGCTGGATTATCAGGA 1049
Qy      426 IleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyr 445
Db      1050 ATCATAATGGGCATAATAATTTGCAAGTGTCTCTCTGTTTATGACACCATTTATTACTGAT 1109
Qy      446 ThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuLeuAspTyr 465
Db      1110 ATACCTCAGTGTGCAATGGCTGCCATTGTGATTTCTGCTGTCATCGGCTGTGATTAAT 1169
Qy      466 GluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAla 485
Db      1170 GNAAGGCGCACTTCTCTGGGGTATTGTAAGAAGGATTTCTTCTGTGGGCGATGACA 1229
Qy      486 TyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSer 505
Db      1230 TTTACTACAACTTAACTTTTGGCATTCAGATTGGTGTCTTGTGGGTCGGGTTTTCG 1289
Qy      506 ValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuLeuGlyAsnIlePro 525
Db      1290 CTGGCAATTTGTGATCCATGAATCTCGAAATCGCATATAGCTGTGTTTGGCGCGTTCCT 1349

```


Db 73 CCACTCGCGCTGGGTACCCGAGCTCGATAGCTCTCCACTACCGCCGCGCTGGTTC 132
 QY 105 LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer 124
 Db 133 CGTCCGGATGTCAGCGCGGCTGTGGTAGCGCAATCCAGATCCCGACCGCATCGCT 192
 QY 125 TyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIlePro 144
 Db 193 TAGCCGACATCGCGCTTCCCGCCAGGTAGGCTCTATGCGTGCATCTCTGCCGATG 252
 QY 145 LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly 164
 Db 253 CTGATCTAGCCCTGATGCGCAGCTCGCGCAGTTGATGTAGGCCCGCCGACGCGCG--- 309
 QY 165 SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspPro-----AsnGluAspPro 182
 Db 310 -----ACCGCGGGATGGTGGCGCGCGCATCACCCGCTGGCGGGCGGCGATCG 360
 QY 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
 Db 361 CAGCGCTGTGGTACCTGTGATGATCGTGGCATCATGCTCGGCTGTCTCCATCGTC 420
 QY 203 LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGly 222
 Db 421 GCGGCTGTGGCGCGCGCGGTTTCATCGCAGCTTCTCTCGCGACCGATCTCTGTGTCGC 480
 QY 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeu 242
 Db 481 TACCTCAAGCGCATCGCGCTGAGCTGTGGTGGCGCAACTGGCGAAGCTGTTCGGCTAC 540
 QY 243 GluHisPheThrHisGly-----AlaAspIleIleSerValMetArgSerValPheThr 260
 Db 541 GAGCGCGCACCGAGCGCTTGTGGCGCGCATCTCGCGCTGTGGAGAACCTGTCTG--- 597
 QY 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280
 Db 598 -----CACATCATTTGGCGCGACGCTGCTCGCGAGCTCTCGCTCTGTGCTGATG 648
 QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
 Db 649 GTGCTCTG-----CCGCGGGCTTC-----CCGAGTGTCCCGGG 684
 QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGlyLysHis 320
 Db 685 GCGCTCTCGCGGCTCTCTCGCGCAGCTCGCGCGCGCTGTGGCGCTGGATCGCTAC 744
 QY 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProSerLeu----- 337
 Db 745 GCGGTGGAACTGCTGGCGAGGTACCGCGCGGCTG-----CCGCAACTGAGCTGGCG 798
 QY 338 ---ThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleVal 356
 Db 799 CAGACCGCTGGAGAACTGAAGACCTGTGCGCGACGCC-----ACCGCATCAG 852
 QY 357 ValGlyIleIleSerLeuAlaGluGlyIleAlaValIleArgSerPheAlaMetTyrLys 376
 Db 853 GTG-----GTCAGTCTTCAGCGCCATGCTCACCGCGCGAGCTTCGCGCGCGCTCAC 906
 QY 377 AsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGly 396
 Db 907 GGCTACAGCATCAACCCCAACCAAGATTCGTGCGCTCGGCTGGCCCAACATCGCGCGC 966
 QY 397 SerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsn 416
 Db 967 GGGGTCTCCAGGGCTTTCGCTCATCAGCGCGCGGCTACCGACCGCGGTGAGCAGCATG 1026
 QY 417 AlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeu 436
 Db 1027 GTCGGGCGGACAGACCAACTGGTGGCGGTGGTGGCGCGCTGTGTATCCCGCCACCTG 1086
 QY 437 LeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleVal 456
 Db 1087 TTGCTGCTGAACAAGCCCTGGGCTGGGTGGCGATGCTCGGCGCTCGGTGCTGCTGTTG 1146

QY 457 SerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLys 476
 Db 1147 CTGCGCGGCTGGGGCTGTATCGAGCTGCAGGCGCTGAAGGGCTTCTGGAGCTCAGCGCG 1206
 QY 477 PheAspPheValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIle 496
 Db 1207 TTTCAGTTTCAGCTCTGCTGCTGACCACTGTCGCGGCTGCTCAGGCTCGCGCTGCG 1266
 QY 497 GlyLeuValIleAlaIleValIleSerValLeuArgValLeuPheIleAlaArgPro 516
 Db 1267 GGAATCTTCGTCGCGCTCAGCATCGCTTACTTCCGCTGCTCTACTACCTATCGCCCG 1326
 QY 517 ArgThrPheValLeuGly-----AsnIleProAsnSerValIleTyrArgAsnVal 533
 Db 1327 AGCAGCGCGGTGCTCGCTGGATGCAGCGCATCGAGCGCGCATGTC-----GAGCTG 1377
 QY 534 GluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaPro 553
 Db 1378 GCCAAGTACCCCGACGCCACCACTCTTCCGCGCTGGTATCTACCGCTTCAGCGCCCG 1437
 QY 554 IleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGlu 573
 Db 1438 CTGCTGTCTTCAAGCGCGAGTACTTCAAGCAGCGGCTGCTGCGGTGTGGAC----- 1491
 QY 574 GluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSer 593
 Db 1492 -----GGCAGCGAGCGCGCAATGCGGTGCTGCTCAACCGCGAA 1530
 QY 594 AlaValGlyAsnIleAspThrSerGlyLeuSerMetLeuGluGluValLysLysIleThr 613
 Db 1531 GCGATGACCAACTCGATCAGATCAGCGGCTGGCCACCTCCAGAGGTGCACAGATCTCTC 1590
 QY 614 GluArgArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLysLysLeu 633
 Db 1591 AAGCGCCAGGGCTGACCTGCTGCGCGGGTGACCGGCGCAGCGCTGACCTGCTGCTG 1650
 QY 634 AsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGlu 653
 Db 1651 CAACGCTCGAGCATCTCGCGGAGATCAAGCGCGCTGGTGTTCAGCTCGGTGCGCTCC 1710
 QY 654 AlaValGlyAlaCysAsnPheAsnLeuArg 663
 Db 1711 GGGGTCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740

RESULT 9

US-09-252-991A-7184
 ; Sequence 7184, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; ORGANISM: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 7184
 ; LENGTH: 1767
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-7184

Alignment Scores:
 Pred. No.: 2,91e-73 Length: 1767
 Score: 676.50 Matches: 172
 Percent Similarity: 50.17% Conservative: 117
 Best Local Similarity: 29.86% Mismatches: 242
 Query Match: 19.47% Indels: 45

DB: 4 Gaps: 13
US-10-762-049-18 (1-680) x US-09-252-991A-7184 (1-1767)
QY 99 TyrThrPheGlnPheLeuLysAlaAepLeuIleAlaGlyIleThrIleAlaSerLeuAla 118
DB 11 TACCGCGCGCTGGTTCGTCGATGTCAGGCGCGGCTGTCGTAGCGCATCCAG 70
QY 119 ileProGlnGlyIleSerTyrAlaLysLeuAlaAenLeuProProIleLeuGlyLeuTyr 138
DB 71 ATCCCCACCGCATCGCTACGCCAGATCGCGGCTTCCGCCCCAGGTAGGCTCTAT 130
QY 139 SerSerPheIleProPheLeuTyrAlaMetMetGlySerSerArgAspLeuAlaVal 158
DB 131 GCGTGATCCTCGCGATGTGATACGCCCTGTACGCGCAGCTCGCGCATGTGATGGTA 190
QY 159 GlyThrValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspPro 178
DB 191 GCGCGCGACGCGCG-----ACCGCGCGATGGTTCGCGCGCCATCACCCCG 238
QY 179 -----AsnGluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAla 196
DB 239 CTGGCGCGCGCGATCCGCGCGCTGGTGGACCTGTGATGATCGTCGATCATGTC 298
QY 197 GlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSer 216
DB 299 GGGCTGTTCTCATCGTCGCGCGCTGGCGCGCGGGTTCATCGCCAGCTTCCTCTCG 358
QY 217 HisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeu 236
DB 359 CGACCGATCCTGGTGGCTTACTCAACGCGATCGCGCTGAGCTGTGTCGGCACTG 418
QY 237 LysSerIleLeuGlyLeuGluHisPheThrHisGly-----AlaAspIleIleSerVal 254
DB 419 GCGAAGCTTCTGGCTACGAGCGCGACGAGCGCTTCGTCGCGGCACTCTCGCGCTG 478
QY 255 MetArgSerValPheThrGlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCys 274
DB 479 CTGGAGAACCTGCTG-----CACATCCATTGGCCGACGCTGATCTCGCGCAGC 526
QY 275 ValPheIlePhePheLeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePhe 294
DB 527 CTCTCGCTTCTGATGTGTGCTG-----CCGCGCGCGCTC 565
QY 295 TrpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPhe 314
DB 566 ---CCGCGATTGCGCGCGCTCTGCGCGCTTCTTCTGCGCAGCTCGCGCGCGCTG 622
QY 315 ThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysGlyLeuAsnPro 334
DB 623 CTGGGCTCGATCGCTACGCGCGTGGAACTGCTTGGCGAGGTACCGCGCGGCTG----- 676
QY 335 ProSerLeu-----ThrAsnLeuValPheValSerProTyrMetThrAla 350
DB 677 CGCAACTAGCTGGCGCGCAGACCGCTGGAGGAACTGAAAGCTGTGCGCGCAGCC 736
QY 351 ValLysThrGlyIleValValGlyIleSerLeuAlaGluGlyIleAlaValGlyArg 370
DB 737 -----ACCGCATCACGCTG-----GTCAGCTTCTGCGCGCATGTCTACCGCGCGC 784
QY 371 SerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGly 390
DB 785 AGCTTCGCGCGCGCTCACGCTACAGCATCAACCCCAACCAAGAAATTCGTCGCCCTCGGC 844
QY 391 ThrMetAsnValValGlySerPheThrCysTyrLeuThrThrGlyProPheSerArg 410
DB 845 CTGGCCCAATCGCGCGCGGCTCTCCAGGGCTTCGCCATCAGCGCGCGCCACTCACGC 904
QY 411 SerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeu 430
DB 905 ACCGCGGTGAACGACATGCTGCGCGCAAGACCAACTGTCGCGGTGTCGCGCGCTG 964
QY 431 AlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValVal 450

DB 965 GTGATCGCGCGCACCTGTTGCTGTAAGCAAGCCCTGGGCTGGGTCCCATGCGCGCG 1024
QY 451 LeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHis 470
DB 1025 CTGCGTGGTCTGTTGCTGGCGGCTGGGCTGATCGACGTCGACGCGTGAAGGCG 1084
QY 471 LeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValVal 490
DB 1085 TTCTGGAAGCTCAGCGCTTCAGTTCAGCTTCGCTGCTGCTGACACCGCTGGGCTGCTC 1144
QY 491 PheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeu 510
DB 1145 AGCGTCGCGTGTGCGCGGAATCTTCGTCGCGCTGACATCGCGTACTGCGCTGCTC 1204
QY 511 LeuPheIleAlaArgProArgThrPheValLeuGly-----AsnIleProAsnSer 527
DB 1205 TACTACACCTATCGCGCGAGCGCGGTGCTCGGCTGGATGCACGCGCATCGACGCCAG 1264
QY 528 VallIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIle 547
DB 1265 GTC-----GAGCTGGCCAAAGTACCCGAGGCCACCACTCTTCCGCGCTGGTGATC 1315
QY 548 LeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThr 567
DB 1316 TACCGCTTCGAGCGCGCTGCTGTTCTTCAACGCGAGTACTTCAAGCAGCGGGTGT 1375
QY 568 ArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyr 587
DB 1376 GCCGTGGTGGAC-----GGCAGCGAGCGCGGAATGCG 1408
QY 588 VallIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607
DB 1409 GTGCTGCTCAACGCGGAAGCCATGACCAACCTCGACATCAGCGCGCTCGCCACCTCCAC 1468
QY 608 GluValLysIleThrArgArgGluLeuGlnLeuValLeuValAsnProValSer 627
DB 1469 GAGTGTCAACAGATCCTCAAGCCGAGCGGTGCACCTGTGCTGGCGGGTGACCGGG 1528
QY 628 GluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIle 647
DB 1529 CAGACGCTGGACCTGCTCAACGCTCGAGCATGCTCGCGGAGATCAAGCGCGCGCTGGTG 1588
QY 648 TyrLeuThrValGluAlaValGlyAlaCysAsnPheAsnLeuArg 663
DB 1589 TTCAGCTCGGTGCTCGCGGTGAGTGCCTATCGCTACTGCTGCTGCGC 1636
RESULT 10
US-09-720-317A-5
; Sequence 5, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-5
Alignment Scores:
Pred. No.: 2,05e-73 Length: 1240
Score: 675.50 Matches: 132
Percent Similarity: 67.22% Conservative: 69

```
Best Local Similarity: 44.15% Mismatches: 97
Query Match: 19.44% Indels: 1
DB: 4 Gaps: 1

US-10-762-049-18 (1-680) x US-09-720-317A-5 (1-1240)

QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
DB 6 AGCCTCAGGAAGCTATCGCGTGTGGCGATCTTCGCTCCGTAAGAGGGGTACAGACTC 65
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
DB 66 GACGGCAACAAGAGAGATGCTGGCCATGGGGTCTCCCAAGTTCCTGCTCTCTGCTCTCG 125
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
DB 126 TGTATGTGGCAACAGGTTCGTCTCCGACCGCAGTGAACCTTCAGCGCGGGGCCAGG 185
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuPheLeuThr 440
DB 186 TCGACCGTTTCAAAACATCGTCATCCACCGTGTTCGTACCCCTGGAGCTGTTTCATG 245
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
DB 246 AAGCTCTCTACTACACGCCCGCTGGCGTCTCGCTCCATCATCTCTGCGGCTCTCCG 305
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
DB 306 GGAATGATGCATCAAGGGCGCTGCAGCATATGAAGATCGCAAGATGGATTTCCTC 365
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
DB 366 ACCTGCTCGGTGGTGTGGCGTCTGTTGGGTGGTGGAGATTGGGCTTCGAGTT 425
QY 501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520
DB 426 GCATCTGGCATTTCTTCGCAAGATCATCATACAGTCGCTCGGCTCAGGTGGAGATC 485
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
DB 486 CTTGGCAGGCTACAAGGACAGATATCTTCGACGCGTCAGGAGTACCCCTGTAGCCTGC 545
QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaPro---IleTyrPheAlaAsnAla 559
DB 546 CTAACCTCCGACTGTACTGCTATACGCGTCGACACATCTCTCTCTCTTCATCAAGCC 605
QY 560 SerTyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAla 579
DB 606 ACTTCGCTCAAGAAAGGATCACAGAGTGGGTTTGGGAAGGAGTGGAGACCTCAAAATGGA 665
QY 580 ThrGlyGluThrSerLeuGlnTyrValIleAspMetSerAlaValGlyAsnIleAsp 599
DB 666 AAGCGAGGAGGAGGATACAGCAGTGTCTTGTATATGTCAAGTGTGTAAACATCGAC 725
QY 600 ThrSerGlyIleSerMetLeuGluGluValLysIleThrGluArgGluLeuGln 619
DB 726 ACTTCAGGACTCACTGCTGGAAGAAATACACAAGGAGTGTGTCTCTCTGCTTACAG 785
QY 620 LeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln 639
DB 786 ATGGCTATAGCCAGTCCGGGATGGAAGGCGATTCAGAAGATGAAGTGTACAGGTGGTG 845
QY 640 AsnHisLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCys 658
DB 846 GACAGGGTAGGACAGGACTGGATCTTCATGACAGTAGGTAGGTGAAGCGGTGGAGCCCTGT 902
```

RESULT 11

```
US-09-785-381-4
; Sequence 4, Application US/09785381
; Patent No. 6602992
; GENERAL INFORMATION:
; APPLICANT: DALLOS, Peter
; APPLICANT: ZHENG, Jing
```

```
; APPLICANT: MADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-37U1
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2441
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-785-381-4

Alignment Scores:
Pred. No.: 7.8e-70 Length: 2441
Score: 651.00 Matches: 191
Percent Similarity: 45.95% Conservative: 144
Best Local Similarity: 26.20% Mismatches: 270
Query Match: 18.73% Indels: 124
DB: 4 Gaps: 21
```

```
US-10-762-049-18 (1-680) x US-09-785-381-4 (1-2441)

QY 35 MetAsnAsnPheGluArgValHisGlnValGluValProPro----- 48
DB 198 CTGCAGACCATGATCATCTCTCAAGAAATGAATCCCTGCAGAGACCCAGAGGTACTAC 257
QY 49 ---ProGlnProPhePhe-----LysSerLeuLysTyrSerLeuLys 61
DB 258 GTGGAAGAGGCCCATCTTCAGTCATCTCTCAAGAGAGCGCTGCACGTCGAAGGACAAA 317
QY 62 GluThrPhePheProAspAspProLeuArgGln---PheLysAsnLysProAlaSerLys 80
DB 318 GTCACAGAGTCCATTGGAGATAAGCTGAAGAGAGGCAATTCACGTGTACTCCT-----AAA 371
QY 81 LysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTyrAlaProLysTyrThr 100
DB 372 AAAATAAGAAACATCATTTACATGTTCTCGCTATCATCAAGTGGTGCACGACATATAA 431
QY 101 Phe---GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIle 119
DB 432 TTCAAGGAGTATGTGTAGTGACTTGGTCTCGGCGATGAAGCACTGGGGGTACTCCAGCTT 491
QY 120 ProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSer 139
DB 492 CCCCAAGGCTTAGCTTCGCCATGTGCGCAGCGCTCCGCTCGCTGTACTCA 551
QY 140 SerPheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGly 159
DB 552 TCGTTTACCCTGTTATCATGTACTGTTCTTTGGNACCTCAAGACACATATCTATAGGT 611
QY 160 ThrValAlaValGlySerLeuLeuMetGlySerMetLeu----- 172
DB 612 CCTTTTGTCTTATTAGCTTGTATGATGTGGAGTGTGGCGTCCGCTCGGTAGTACAGATGAT 671
QY 173 -----SerAsnAlaValAspProAsnGlu---AspProLysLeuTyr 185
DB 672 ATTGTATCCAGGAGGAGTAAATGCAACCAACGGGACAGCCAGAGATGCACCTAAGA 731
QY 186 LeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeu 205
DB 732 GTGAAGTGGCCATGCTGTGTACCTTACTTTCAGGAATCATTCAGTTTGGCTTAGGTGTC 791
QY 206 PheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGly 225
DB 792 TGTAGGTTGGATTGTGGCCATATACCTCACGGAGCCATTGGTGGAGGCTTTTACCAC 851
QY 226 GlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPhe 245
DB 852 GCGGCTGTGTCCAGCTGTTCATCCATGCTTAAATAATACCTGTTTGGGGTCAAAACAAAG 911
```

Qy	246	ThrHisGlyAlaAspIleIleSerValMetArgSer-----ValPheThrGlnThr	262
Db	912	CGGTAC---AGTGGAAATCTTTTCAGTGGTGTATAGTACAGTGTGTTGTGTCAGAAATGTT	968
Qy	263	HisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSer	282
Db	969	AAAAACCTCAACGTGTGTTCCCTAGGCGTGGCGCTGATGGTTTTTGGTTGGTGTGTTGGGT	1028
Qy	283	ThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeu	302
Db	1029	GGCAAGAAATTTAATGAGACA-----TTTAAAGAGAAATGCGCAGCACCCCAT	1076
Qy	303	Thr-----SerValIleLeuGlySer---LeuLeuValTyrPheThrHisAla	317
Db	1077	CCTCTAGAGTCTTTGTGTGGTGAATGGGACTGGCATTTCTGCAGGATTTAAACCTACAT	1136
Qy	318	GluLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProPro-----	335
Db	1137	GAGTCCTACAGTGTGGATGTCGTGTGGAACACTTCTCTCGGGGTACTTCTCTCGGGCAAC	1196
Qy	336	-----SerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLys	352
Db	1197	CCAGACACACGCTGTCACCTGGTGTATGTGGAC-----	1232
Qy	353	ThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPhe	372
Db	1233	---GCCATTGCCATCGCATCGTTGGATTTTCAGTGCAGATCTCCATGGGCCAAAACCTTG	1289
Qy	373	AlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMet	392
Db	1290	GCAAAATAAGCATGGCTACCAAGGTGTGATGGCAATCAGGAGCTCAATGCGTTGGGGATATGC	1349
Qy	393	AsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAla	412
Db	1350	AATCCATTGGAICTCTCTTCCAAACCTTCTCGATTTCCGTCTCTGCTCGAGGCTT	1409
Qy	413	ValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaVal	432
Db	1410	GTTCCAGGAAGAACTGGAGGGAAAAACACAGCTTCGAGGTGTTGTTGGCTCTGTTGATGATT	1469
Qy	433	MetLeuThrLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSer	452
Db	1470	CTGTGTGTCATATTAGCCACCGGATTCCTCTTTGAGTCGTTACCCACAGGCTGCTCTTCC	1529
Qy	453	AlaIleValSerAlaMetLeuGlyLeu---IleAspTyrGluAlaAlaIleHisLeu	471
Db	1530	GCAATTGTGATGTCGAACCTGAAGGAATGTTATGCAAGTCTTCAGAGTCTCAGACCTGCCTTTTTT	1589
Qy	472	PheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPhe	491
Db	1590	TGGAGAACCAAGAAAAATAGAGCTGACCATCTGGCTGACCACTTTGTGTCTCTCCCTGCTC	1649
Qy	492	GlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeu	511
Db	1650	CTCGGCTTGGACTACGCACTGATTACCGCGCTGATCATGCTCTGCTTCACAGTGATTTAT	1709
Qy	512	PheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArg	531
Db	1710	AGAACACAGAGTCCCAAGCTCAAAAGTCTCTGGGGCAGCTCTCCCTGACCGGATGTGTACATT	1769
Qy	532	AsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAsp	551
Db	1770	GACATAGATGCAATGAGGAGGTGAAGAAATTCCTGGAATAAAAATATTTCCAAATAAAT	1829
Qy	552	AlaProIleTyrPheAlaAsnAla-----	559
Db	1830	GCCCCAATTTACTATGCAAAATAGCGACTTGTATAGCAGCGCTTTAAAAAGAAAGACTGGA	1889
Qy	560	-----	564
Db	1890	GTAACCCAGCACTCATATTATGGGAGCGAGAAAGAGCCATGAGGAAGTACGCCACAGAA	1949
Qy	565	-----ArgIleThrArgTrpIleAsp---	571

```

Db      1950 GTTGGAAATGCCAACGTGCGCAATGCTACTGTGTTGCACAAAGTAGCATCCAGAAGTAGCAGGA   2009
Qy      572 -----GlUGluGluGluArg-----|||||:::|||||||-----ile 577
Db      2010 GAAATGCTACAAAACTGGAAGAAGAGGATGATGAAGTCAAAATTCCCCCAATAGTCATC   2069
Qy      578 LyAlaThrGlyGluThrSerLeuGln-----|||||:::|||||||-----Tyr 587
Db      2070 AAAACAACAATTTCTCTGAAGAGCTGCAGAGATTTTTGCCCGGCGGAAAAAATGTCCACACT   2129
Qy      588 ValIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu   607
Db      2130 GTCATCTAGACTTTTGACAGGTCAAATTTTGTGGATCTGTGTGGATGTAATACTCTGGCC   2189
Qy      608 GluValLysLysIleThrGluArgAgGluLeuGlnLeuValLeuValAsnProValSer   627
Db      2190 GGGATTCTGAAAGAATATCGAGATGTTGGAATTTATGTATATTTAGCAGGATGCAGCCCA   2249
Qy      628 GluValMetLysLysLeuAsnLysSerLys---PheGlnAsnHisLeuGlyLysLysTrp   646
Db      2250 CAAGTTGTGAATGACCTCACCCGCAACAACCTTTTGAANAATCTGCCTTTGAAAAGAGCTT   2309
Qy      647 IleYrLeuThrValGluGluAlaVal   655
Db      2310 CTGTTCCAGTATCCAGATGCAGTC   2336

RESULT 12
US-09-785-381-2
; Sequence 2, Application US/09785381
; Patent No. 6502992
; GENERAL INFORMATION:
; APPLICANT: DALLOS, Peter
; APPLICANT: ZHENG, Jing
; APPLICANT: MADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-37UL
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 4113
; TYPE: DNA
; ORGANISM: Meriones unguiculatus
US-09-785-381-2

Alignment Scores:
Pred.No.:          2,66e-69           Length:    4113
Score:             650.00            Matches:    189
Percent Similarity: 46.63%           Conservative: 143
Best Local Similarity: 26.54%        Mismatches:  254
Query Match:       18.71%            Indels:     126
DB:                4                 Gaps:       21

US-10-762-049-18 (1-680) x US-09-785-381-2 (1-4113)

Qy      39  GluArgValHisGlnValGluValProProGlnProPhePhelysSerLeuLysTyr   58
Db      311 GAGAGGCTGCAT-----|||||:::|||||||-----GTCAAGGAC   331
Qy      59  SerLeuLysGluThrPhePheProAspAspProLeuArgGln---PheLysAsnLysPro   77
Db      332 AAAGTCTCAGAGCTCAATT-----GGGGATAAGCTGAAGCAGCGGTTCACATGCACCTCCC   385
Qy      78  AlaSerLysLysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaPro   97
Db      386 -----AAAAGATAAGAAACATCATTTACATGTTCTGCGCCATCACTAAGTGGTTGCCCA   439
Qy      98  LysTyrThrPhe---GlnPheLeuLysAlaAspIleuIleAlaGlyIleThrIleAlaSer   116

```

Db 440 GCTTACAAGTCAAGGAGTATGTGTTGGGTGACTTGTTTTCAGGCATAAGCACCAGCGGTG 499
QY LeuAlaIleProGlnGlyLeuSerTyrAlaLysLeuAlaAsnLeuProPheLeuGly 136
Db 500 CTTTCAGTTCCTCCAGAGCTTGGCTTGCATGCTCGGGCTGTTCTCGGTGTTCCGGC 559
QY LeuTyrSerSerPheLeuProProLeuLeuTyrAlaMetMetGlySerSerArgAspLeu 156
Db 560 CTGTACTCTTCATTTTATCTGTATCATGTACTGTCTTTCTTTGGGACCTCCAGACACATA 619
QY 157 AlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMetLeu----- 172
Db 620 TCTATAGGTCCTTTTCGCGGTTCATTAGTTCGATCGGTGCTGCTGCGGTGCTG 679
QY 173 -----SerAsnAlaValAspProAsnGlu---AspPro 182
Db 680 CCGATGACATCGTCATCCCGGAGGAGTGAACGCAACCGCAGCGAGGCCGAGAC 739
QY 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLa 202
Db 740 CGCTGAGAGTGAAGTGCAGTGTCTGCACCTGCTCTCAGGAATCATTCAGTTTGC 799
QY 203 LeuGlyLeuPheArgLeuGlyLeuLeuValAspPheLeuSerHisAlaThrIleLeuGly 222
Db 800 CTAGGTGTGTCAGGTTTGGATTGTGGCCATATACCTCACGGAGCGCTGCTGCGAGGG 859
QY 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeuGlyLeu 242
Db 860 TTCAACACCGCCCGCGGTGACGCTTCACATCCATGTTGAATACCTGTTTGGGTT 919
QY 243 GluHisPheThrHisGlyAlaAspIleIleSerValMetArgSer-----ValPhe 259
Db 920 AAGACAAAGCGGTAC--AGTGGATCTTTTCGGTGTATATAGTACAGTTGCTGTG 976
QY 260 ThrGlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePhe 279
Db 977 CAGAATGTTAAAAACCTCAACGTGTGTTCTTCCTAGGCGTGGCTGTATGTTTGGTTTG 1036
QY 280 LeuLeuSerThrArgTyrPheSerLysLeuArgProArgPhePheTrpValSerAlaMet 299
Db 1037 CTGTGGTGGCAAGAGGTTTAATGAGAGA-----TTTAAAGAGAAATGGCCA 1084
QY 300 AlaProLeuThr-----SerValIleLeuGlySer---LeuLeuValTyrPhe 314
Db 1085 GCACCATTCCTCTAGAGTCTTTTGTGTGTGTGATGGAACTGGCATTTCCGGGGTT 1144
QY 315 ThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnPro 334
Db 1145 AACTTGCACGATCTCTACAGTGTGGATGTGTTGGAACTCTTCTCTGGGGCTACTCCCT 1204
QY 335 Pro-----SerLeuThrAsnLeuValPheValSerProTyrMetThrThr 349
Db 1205 CTTGCCAACCCGGACACCGCTCTTCCACCTCGTGTATGTGGAT----- 1249
QY 350 AlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGly 369
Db 1250 -----CCCATGGCATGCCATCGTTGGATTTCAGTGACAATTTCCATGGCC 1297
QY 370 ArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIle 389
Db 1298 AAAACCTTGGCGAATAAGCATGTGCTACCGAGTTGATGGCAATCAGGAGCTCATCGCTTG 1357
QY 390 GlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSer 409
Db 1358 GGGATATGCAACTCCATCGGATCTCTTCCAGACCTTCTCCATTTCCCTGCTCTGTCT 1417
QY 410 ArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSer 429
Db 1418 CGAGGCTGTTTCAGGAGGGAACCTGGAGGGAACACAGCTCGCAGGTTGCTTGGCGCTG 1477
QY 430 LeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuVal 449
Db 1478 CTGATGATTCTGCTGCTGTTTATAGCCACTGGATTCTCTTTGAGTCAITTCGCCCGAGGCT 1537

QY 450 ValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu---IleAspTyrGluAlaLa 468
Db 1538 GTGCTCTCGCCATTTGATCGTGAACCTGAAAGGATGTTTATGTCAGTCTCAGATCTG 1597
QY 469 IleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGly 488
Db 1598 CCCTCTCTTCTGGAGAACCAACGACCAATAGAGCTGACCATCTGGCTTACCACCTTTGTGTC 1657
QY 489 ValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArg 508
Db 1658 TCCTCTGTTCTCGGCTTGGACTACGAGCTGATTACTGCTGTGATCAATCTCTGCTGACT 1717
QY 509 ValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerVal 528
Db 1718 GTGATTACAGAACCCAGAGTCGAGCTACAGGTCTCTGGGCGAGCTCCTGACACCGAT 1777
QY 529 IleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeu 548
Db 1778 GTATACATTGACATGACGCATATGAGGAGGTGAAGAAATTCCTGGAATAAAATATTC 1837
QY 549 GluIleAspAlaProIleTyrPheAlaAlaAla----- 559
Db 1838 CAGATAAACGCCCAATTACTATGCAACAGTACTTGTATAGCAACGCCCTAAAGA 1897
QY 560 -----SerTyr 561
Db 1898 AAGACTGGTGTGAACCCAGCGCTCATATGGGAGCAAGGAGGCCATGAGGAAGTAC 1957
QY 562 LeuArgGlu-----ArgIleThrArgTrp 569
Db 1958 GCAAAGGAAGTCGAAACGCCCAACATTCGCAACGCGAGCTGTTGTCAAAGTGGATGAGAA 2017
QY 570 IleAsp-----GluGluGluGluArg----- 576
Db 2018 GTAGATGGAGAAATGCTACGAAGCCCGAAGAGAGGATGATGAAGTAAATATCCCCCA 2077
QY 577 -----IleLysAlaThrGlyGluThrSerLeuGln----- 586
Db 2078 ATAGTATCAACAAACAATTTCTCGAGAGCTGCAGAGATTATGCCCCAGACAGAAAT 2137
QY 587 -----TyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSer 604
Db 2138 GTCCACACTCATTTCTAGACTTCACAAAGTCAATTTTATCGACTCTCTGGAGTAAAA 2197
QY 605 MetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsn 624
Db 2198 ACCCTGGCTGTGATGGTGAAGAAATACCGAGATGTTGGTATTATGTACTTAGCAGGA 2257
QY 625 ProValSerGluValMetLysLysLeuAsnLysSerLys---PheGlnAsnHisLeuGly 643
Db 2258 TCAGCCCAAGTGTGATGATGACCTCACCCGCAACCGTTTCTTTGAAATCTGCTTA 2317
QY 644 LysLysTrpIleTyrLeuThrValGluGluAlaVal 655
Db 2318 AAAGAGCTTCTGTTCCACAGTATCCATGATGTCAGTC 2353

RESULT 13

US-09-902-540-7746
; Sequence 7746, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825

```
; SEQ ID NO 7746
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-7746

Alignment Scores:
  Pred. No.:      3,76e-69      Length:      1773
  Score:          643.50       Matches:      162
  Percent Similarity: 50.44%   Conservative: 124
  Best Local Similarity: 28.57% Mismatches:    256
  Query Match:      18.52%    Indels:         25
  DB:                4        Gaps:          7

US-10-762-049-18 (1-680) x US-09-902-540-7746 (1-1773)

QY 91  ProIlePheGluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAspLeuIleAla 110
Db 67  CCGGGCTCGGGCAGGACCGCGGTATCAGCGCGCTGTGTTCCGGCGGACCTGCTCTCG 126
QY 111 GlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAlaAsn 130
Db 127 GCGCTCACCATTGGTCGATGCTCATCCCGAGGGCTCGCTACGCGCAGCTCGTGGGC 186
QY 131 LeuProIleLeuGlyLeuTyrSerSerPheIleProProIleTyrIleAlaMetMet 150
Db 187 GTGCGTCCAGCAGCGGCGCTCTACGCGGCGGTGGTGGGAATGCTGGCCATCGCTGTTTC 246
QY 151 GlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySer 170
Db 247 GGGCCCTCGCGGACCTCATCATCGGACCGGAGCGGGTGGCCGCACTCATCCCGCGG 306
QY 171 MetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAlaPhe 190
Db 307 GCGCTCGCG--CCGGTGGCGGGCGGCGGCGGCGGCTTACGATCGCTGGCGGCA 363
QY 191 ThrAlaThrIleuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGlyLeu 210
Db 364 CTGCTGGCGCTGTGCTCGCGGTGCTGAGCCTGCTCGGAGGGCTGCTCAAGTGGCGCG 423
QY 211 IleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrVal 230
Db 424 CTGCGGAGATTCTGTGCCAGCCCATCTCATCGGCTATCATCATGTCGCGCGGCTCATC 483
QY 231 ValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis-----PheThrHisGly 248
Db 484 ATCATCGGACGCGAGCTCGCGGCTCTTCGGGAGGAGCGGCGGCTCGGACAGCTTCTCG 543
QY 249 AlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGlu 268
Db 544 GGTTCAGGTGTTTCAGGTGGCCACCCACCTC--GAACGAGACCCACGTC-----CCG 591
QY 269 SerAlaValIleGlyCysValPheIlePheLeuLeuSerThrArgTyrPheSerLys 288
Db 592 ACCTCTCTGCTGGGGCTGGGGCTCATCAGCGGCTGCTGCTGCTGGGCACTCTCGGCC 651
QY 289 LysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGly 308
Db 652 AAGGTGCC-----GGTCCGCTCATCTCGTGGTGGTCTCTCACC 687
QY 309 SerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeu 328
Db 688 ACGGTGGCGGAGGGCTGTTCAGTTGGAGCATGGGGGTCATCAAGGTCTGGGCGCCCTTC 747
QY 329 LysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThr 348
Db 748 GCGCGCGAGCCCCCTCCAGGCTTCGCTCCCTCGGCTTCGAGGAC-----GTGCGG 801
QY 349 ThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaVal 368
Db 802 TCGCTGCTCCCGCGGCTTCAGCCTTCAGCCTGGGCTCGTCAACTACGCCAGCTCGGTGAGC 861
QY 369 GlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAla 388
Db 862 GCGCGATTCTACGCGCAGGAGTTCGCTACCGGTGGACAGCACACGAGGAGTTCCTCGGG 921
QY 389 IleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPhe 408
Db 922 CAGCGCGCAGCAACCTCGCCTTACGCGAGGGGTTCCTCGGTGACGCGGAGCGAC 981
QY 409 SerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMet 428
Db 982 TCGCGCAGCGCGCTCAACGCTCTCCATGGGGGGGGAGACCCAGCTCGTGGGGGTGCTCGCC 1041
QY 429 SerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeu 448
Db 1042 GCAGGGGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1101
QY 449 ValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAla 468
Db 1102 GTGACGCTGGCGCCATCGTCTTCGTGCGCGGTGTACCTGCTGGAGTTCGAGCCATC 1161
QY 469 IleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGly 488
Db 1162 ATCGACCTGTGGCGGTGCGCGGTGAGCGGTGCTCGCGTGGCGGTGCGGTGCGGTGCGG 1221
QY 489 ValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArg 508
Db 1222 GTGCTGTGTGGGATCTCTCAAGGATCTCTGCTCGCGGTGCGGTGCGGTGCGGTGCGG 1281
QY 509 ValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerVal 528
Db 1282 CTCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1341
QY 529 IleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeu 548
Db 1342 GGTCTACCAACGATCGAGCGGTTCGAGAACCGCGAGCGGTGCGCGGTGCGGTGCGGTGCG 1401
QY 549 GluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArg 568
Db 1402 CGCTTCGATGCGCGCTGCTCTTCGCCAACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1466
QY 569 TrpIleAspGluGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrVal 588
Db 1447 -----GAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
QY 589 IleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 608
Db 1498 GTGCTGTATGCGTCCGCGCTGTCGACATGGAGCGTCACCGCTGCGCGGTGCGGTGCGG 1557
QY 609 ValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValLeuValLeuValSer 628
Db 1558 CTGCGCGGTGAGTTCGAGGAGCGCGGTGCTGCTGGCGGTGCGCGGTGCGCGGTGCGG 1617
QY 629 ValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIleTyr 648
Db 1618 CTGCGCGCGCTGCTCGCGGTACCGCGCTGCTGCGAGCGGTGCGCGGTGCGCGGTGCG 1677
QY 649 LeuThrValGluAlaVal 655
Db 1678 GCCACCGTGGAGCGCGGTG 1698

RESULT 14
; Sequence 768, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
```

```

; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 768
; LENGTH: 4854
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902,540-768

```

Alignment Scores:

Pred. No.:	2.31e-68	Length:	4854
Score:	643.50	Matches:	162
Percent Similarity:	50.4%	Conservative:	124
Best Local Similarity:	28.5%	Mismatches:	236
Query Match:	18.5%	Indels:	25
DB:	4	Gaps:	7

US-10-762-049-18 (1-680) x US-09-902-540-768 (1-4854)

[illegible]

RESULT 15

```

RESUB 15
US-09-949-016-4516
; Sequence 4516, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```

```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4516
; LENGTH: 2832
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4516

```

Alignment Scores:	
Pred. No.:	1,558-68
Score:	641.50
Percent Similarity:	44.9%
Best Local Similarity:	25.27%
Query Match:	18.46%
DB:	4
Length:	2832
Matches:	184
Conservative:	143
Mismatches:	290
Indels:	111
Gaps:	22

US-10-762-049-18 (1-680) X US-09-949-016-4516 (1-2832)

Qy	18	AenThrIleileGluAspMetGlySerValAspTyrGluTyrProLeuGlyMetAen---	36
Db	55	AAGTTTTACCCAGAGACTCAGCTGAAGAAATGACAGTTATCCATCTGGGATCCATCTG	114
Qy	37	-----AenPheGluArgValHisGlnValGlu	45
Db	115	GAACTTCAAAGGGAATCAAGTACTGACTTTCAGCAATTTGAGACCAATGATCAATGC-	171
Qy	46	ValProProGlnProPhePheLysSerLeu-----LysTyrSerLeuLysGluThr	63
Db	172	-----AGACCTTATCATGAGTACCTTATTGAGCGTCAAGAGAATCAGATA	219
Qy	64	PhePheProAspProLeuArgGlnPhe--LysAen-----LysProAla---	78
Db	220	AACTTCAAGAGTTGTATTAAAAGCTGAGAGAATGCCAGTCGAGTCCAGCCAAA	279
Qy	79	SerLysLysPheMetGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLys	98
Db	280	GCCAAAAATATGATTTTAGT-----TTCTCTCTGTTTGCGAGTGGCTCCCAAAA	330
Qy	99	TyrThrPhe---GlnPheLeuLysAlaAspLeulealaglyllethrIleAlaSerLeu	117
Db	331	TACGACCTAAAGAAAAAACATTTTAGGGGATGTGATGCAGGCTTGATTTGGGGCATATTA	390
Qy	118	AlaIleProGlnGlySerTyralalylsleualaenLeuProIleLeuGlyLeu	137
Db	391	TTGGTCCCCAGTCGATTCCTTATTCCTGTGCTGGCCGAAGAACTGCTATAGTCTG	450
Qy	138	TyrSerSerPheIleProLeuIleTyraAlaMetMetGlySerSerArgAspLeuAla	157
Db	451	TACACATCTTTTTTGGCCAGCATCATTTATTTCTCTGGGTACTCCCGTCACATCTCT	510
Qy	158	ValGlyThrValalaValGlySerLeuLeuMetGly-----	169
Db	511	GTGGGCATTTTGGAGTACTGTCGCTTATGATTGCTGAGACAGTTGACCAGAACTACAG	570
Qy	169	-----	169
Db	571	AAAGCTGGCTATCACAAATGCCCATAGTCTCTCTCTTAGGAATGGTTTCAAATGGGAGC	630
Qy	170	SerMetSerSerAsnAlavalAspProAsnGluAspProLysLeuTyrr---LeuHisLeu	188
Db	631	ACATTTAAATACATACATCAGCAGGATATGTGACAAAAGTTGCTATGCAATTTATGGTT	690
Qy	189	AlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuGlyLeuPheArgLeu	208
Db	691	GGCAGCACTGTAACTTTATAGCTGGAGTTTATCAGGTAGCGATGGGCTCTTTCAAGTC	750
Qy	209	GlyLeuIleValAspPheLeuSerHisAlaThrIleleGlyPheMetGlyValalaLa	228

751	GGTTTGTGTTCTCTACCTCTCAGATGCTTCTGCTGAGTGGATTGTTGCTACTGCTGGCTCC	810
229	ThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGlu---HisPheThrHis	247
811	TTCACTATTCTTACATCTCAGGCCAAGTATCTTCTTTGGGCTCAACCTTCCTCGACTAAT	870
248	Gly---AlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArg	266
871	GGTGTGGGCTCACTCATCACTACCTCGATACATGTCTTCAGAAACATCCATATAGACCAAT	930
267	TrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgz-----	284
931	CTCTGTGATCTTAACACCGCCTTTTGTGCTTTTGGTCTTTTGGTTCCTTGGCAACCAAGAACTC	990
285	-----TyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaPro---	301
991	AATGAACACATTCAAATCCAG-----CTTAAGGCACCGATTCTCTATT	1032
302	---LeuThrSerValIleLeuGlySerLeuLeuValTyrPhe-----ThrHisAlaGlu	318
1033	GAACTTGTGTTGTTGTAGCAGCCACATTAAGCTCTCATTTTGGAAACATACAT---GAA	1089
319	LysHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThr	338
1090	AAITATAATTCTAGTATGTCTGGACATATATCCCATCGGGTTATGTCACCCCAAGTACCA	1149
339	-----AsnLeuValPheValSerProTyrMetThrThraValLysThrGlyIleVal	356
1150	GAATGGAACTTAATT-----CCTAGTGTGCTGTAGATGCAATAGCT	1191
357	ValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLys	376
1192	ATTTCATCATCTGGTTTGTATCACTGTATCACTTTCTGAGATGTTTGGCAAGAAACAT	1251
377	AsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGly	396
1252	GGTTACACAGTCAAAAGCAACCAAGGAAATGTATGCCATTGGCTTTGTATAATATCACTCC	1311
397	SerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsn	416
1312	TCCTTCTTCACCTGTTTACTACTAGTCGACGCTCTTGCAAGACATTTGGTTAAAGAATCA	1371
417	AlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeu	436
1372	ACAGGCTGCCATCACTCAGCTTTCTGGTGTGTGAACAGCCCTGCTCTTTTGTGGTCTC	1431
437	LeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleVal	456
1432	CTAGTATAGTCTCTTTGTTCTATTCCTTCAAAAGTGTCTCTGGTGTGATCACAAT	1491
457	SerAlaMetLeuGly---LeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAsp	475
1492	GTAATCTACGGGAGCCCTTCGTAATTTAGGAGATCTCCCAAAATGTGGAGTATTAGT	1551
476	LysPheAspPheValValCysMetSerAlaTyrIleGlyValValPheGlySerValGlu	495
1552	AGAATGGATACATGTTATCTGGTTTGTGTACTATGCTGCTCTGCTGCTGCTAAGTACTGAA	1611
496	IleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArg	515
1612	ATAGGCCACTCTGTGGGGTTGTTTCTATATTTGTGTGATCTTCCTCGCATCAGAAG	1671
516	ProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHis	535
1672	CCAAAGATTCACTGCTTGGCTGGTGGGAAGAGTCTGAGGTCTTTGNAATCTGTGCTGCT	1731
536	TyrGlnAsnAlaLysHisValProGlyMetLeuLeuGluIleAspAlaProIleTyr	555
1732	TACAAGAACCCTTCAGACTAAGCCAGGCATCAAGATTTCGCGCTTGTAGGCCCTCTCTAC	1791
556	PheAlaAsnAlaSerTyrLeuArgGluAArgIleThrArg-----	568
1792	TACATAAACAAGAAGTCTTTAAATCTGCTTTTATACAAACAAACACTGTCAACCCCAATCTTA	1851

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 12, 2005, 18:50:46 ; Search time 996 Seconds
(without alignments)
4483.658 Million cell updates/sec

Title: US-10-762-049-18
Perfect score: 3475
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGWNV 680

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Q=/cgn2_1/USPTO_spool_p/US10762049/runat_12092005_120742_7258/app_query.fasta_1.839
-DB=Published Applications NA -QFMT=fastcap -SURFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -LOCAL=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10762049 @CGN 1 1 723 @runat_12092005_120742_7258
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5
-FGAPO=6 -FGAPEXT=7 -YGAPO=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2595.5	74.7	1977	9	US-09-938-842A-1502 Sequence 1502, Ap
2	2595.5	74.7	1977	11	US-09-938-842A-1502 Sequence 1502, Ap
3	2330.5	67.1	2372	19	US-10-437-963-38184 Sequence 38184, A
4	2036	58.6	2008	18	US-10-424-599-131116 Sequence 131116, A
5	1856	53.4	2727	20	US-10-425-115-50601 Sequence 50601, A
6	1845.5	53.1	2651	20	US-10-425-115-182207 Sequence 182207, A
7	1844.5	53.1	2366	18	US-10-424-599-651203 Sequence 651203, A
8	1811.5	52.1	2001	19	US-10-437-963-95853 Sequence 95853, A
9	1773.5	51.0	2780	19	US-10-437-963-77005 Sequence 77005, A
10	1762	50.7	2656	18	US-10-425-114-15719 Sequence 15719, A
11	1762	50.7	2713	20	US-10-425-115-75763 Sequence 75763, A
12	1734.5	49.9	2429	19	US-10-437-963-73410 Sequence 73410, A
13	1698	48.9	4390	19	US-10-437-963-77012 Sequence 77012, A
14	1676.5	48.2	1518	18	US-10-424-599-131117 Sequence 131117, A
15	1563.5	45.0	2251	19	US-10-437-963-72062 Sequence 72062, A
16	1513.5	43.6	1980	17	US-10-260-238-1066 Sequence 1066, Ap
17	1500	43.2	2336	18	US-10-424-599-52175 Sequence 52175, A
18	1460	42.0	1992	18	US-10-425-114-13333 Sequence 13333, A
19	1446.5	41.6	3252	19	US-10-437-963-78357 Sequence 78357, A
20	1436	41.3	3201	20	US-10-425-115-44188 Sequence 44188, A
21	1406	40.5	2242	20	US-10-425-115-40196 Sequence 40196, A
22	1379.5	39.7	2506	18	US-10-424-599-74393 Sequence 74393, A
23	1218	35.1	2156	20	US-10-424-599-60281 Sequence 60281, A
24	1215	35.0	2156	20	US-10-739-930-3837 Sequence 3837, Ap
25	1204	34.6	1809	18	US-10-424-599-83137 Sequence 83137, A
26	1185.5	34.1	1632	18	US-10-424-599-74389 Sequence 74389, A
27	1175.5	33.8	2313	19	US-10-437-963-76453 Sequence 76453, A
28	1134.5	32.6	2346	20	US-10-425-115-22539 Sequence 22539, A
29	1016	29.2	1303	18	US-10-424-599-139734 Sequence 139734, A
30	1010.5	29.1	1551	18	US-10-424-599-54350 Sequence 54350, A
31	1009	29.0	1247	18	US-10-424-599-130951 Sequence 130951, A
32	1002	28.8	2058	9	US-09-938-842A-717 Sequence 717, App
33	1002	28.8	2058	11	US-09-938-842A-717 Sequence 717, App
34	955.5	27.5	2388	19	US-10-437-963-92700 Sequence 92700, A
35	934	26.9	680	18	US-10-424-599-32307 Sequence 32307, A
36	923	26.6	1019	19	US-10-767-701-11661 Sequence 11661, A
37	901	25.9	1174	20	US-10-425-115-75765 Sequence 75765, A
38	900	25.9	536	18	US-10-424-599-32308 Sequence 32308, A
39	898	25.8	674	19	US-10-767-701-701 Sequence 701, App
40	892.5	25.7	1115	19	US-10-767-701-9435 Sequence 9435, Ap
41	887	25.5	1069	20	US-10-425-115-3619 Sequence 3619, Ap
42	814	23.4	1802	19	US-10-437-963-7407 Sequence 7407, Ap
43	737.5	21.2	1018	18	US-10-424-599-18417 Sequence 18417, A
44	736.5	21.2	841	18	US-10-424-599-135930 Sequence 135930, A
45	715	20.6	788	20	US-10-425-115-75767 Sequence 75767, A

ALIGNMENTS

RESULT 1
US-09-938-842A-1502
; Sequence 1502, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krebs, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24

```

; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1502
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1502

Alignment Scores:
Pred. No.: 5,93e-263 Length: 1977
Score: 2595.50 Matches: 498
Percent Similarity: 87.73% Conservative: 81
Best Local Similarity: 75.45% Mismatches: 77
Query Match: 74.69% Indels: 5
DB: 9 Gaps: 3

US-10-762-049-18 (1-680) x US-09-938-842A-1502 (1-1977)

QY 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
DB 1 ATGGGACCGAGGACTACATCTCCCTCAAGGACGGAGGAGCTTCACCGCGCCACCCAC 60

QY 44 ---ValGluValProProGlnProPhePheLeuSerLeuLysTyrSerLeuLysGlu 62
DB 61 ACGGTGGAGGCTCCACAACTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120

QY 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
DB 121 ACTCTGTTTCCAGACGACCTTTTAGACAAATTTAAGAACCAAAATGTCATCAAGAAATTT 180

QY 83 MetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLysTyrThrPheGln 102
DB 181 GGTGTTAGGCTCAATACTCTCCCGCATTTTCGAATGGGACCCAGCTCAATCTCAAG 240

QY 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
DB 241 TTCTTCAATCAGATCTCATCGCGGAAATCAGCATCGTAGTCTCGCCATCGCTCAGGCG 300

QY 123 IleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerPheIle 142
DB 301 ATCAGTTACGCCAAACTTGCTAACTTGCCCCCAATTTCTTGGGCTTTATTCGAGTTTGA 360

QY 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
DB 361 CCGCCATTGGTATACCGGTGCTAGGAGTTCAAGGAGCTTAGCGGTGGACGGTTGGG 420

QY 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
DB 421 GTTGGCTCTCTGTGACAGGTGGATGCTGAGCAAAAGAGTTGATGCTGAGAAAGATCCT 480

QY 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
DB 481 AAGCTTTTACCTTCACCTTGTTCACCGGCACATTTTTCGCGCGGCTTCTCAAGGCTCT 540

QY 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGln 222
DB 541 CTTGGAAATTTTCAAGGTTAGGTTTCAATAGGACTTCTATCGCATGCAACGATAGTAGG 600

QY 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242
DB 601 ATTCATGGGAGGACGACGCGTGTGAGTCTGCAACAGCTTAAGGGTATTTTCGGACT 660

QY 242 uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnh 262
DB 661 TAAACATTTTCAACAGACTCTACCGATGTTATCTCTGTCATGCGGTTCTCCCAAAC 720

QY 262 rHisGluTyrArgTyrPgluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
DB 721 TCAGAGTGGAGATGGGAAGTGGCGTTCTTGGATGTGGTTTCTTATTTCTTCTCTC 780

```

```

QY 282 rThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLe 302
DB 781 CACCAGATATTTGAGCATCAAGAAACCAAAATCTTTTGGGTGGCGGATGGCTCCTTT 840

QY 302 uThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVa 322
DB 841 GACCTCAGTGATTTCTTGGAACTCTTGGTATTCTTCACTCAGCTCAGGACATGGTGT 900

QY 322 lGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342
DB 901 TCAAGTGATGGGACCTGAAAGAGGTTTGAATCCACTCTCCGTTCTGATCTCATCT 960

QY 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLe 362
DB 961 TACTTCCCTTACATGTCACAGCTGTCACAACTGGCTCATCTCATCTGTCATCTGCTCT 1020

QY 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGln 382
DB 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTTACACATAGACGG 1079

QY 382 yAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy 402
DB 1080 GAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCCTTTCACATCTT 1139

QY 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422
DB 1140 CCTCAACACCGGACCATTTTCAAGTCCGCGAGTGAACATCAACCGGGGTTGCAAGACCGC 1199

QY 422 aAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLe 442
DB 1200 AATGTCCAAATAGTAGTGGGATTCGCTTATGTTTCACTCTCTCTCTCTCTCACACCGCT 1259

QY 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462
DB 1260 TTTTTCACATACACACCTCGTCTCTCTGCCATCATCATATCCGCAATGCTCGGACT 1319

QY 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482
DB 1320 CATTGACTATCAAGCTGCCATCTCTGAAAGTTGCAAGTTTCGACTTCCTCGTCTG 1379

QY 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI 502
DB 1380 CATGAGCCCTACGTTGGGTCGTATTCGCGAGTGTAGAGATTGAGTTCGCTGAGCGGT 1439

QY 502 eValIleSerValLeuLeuValLeuPheIleAlaArgProArgThrPheValLeuGln 522
DB 1440 GCGCATATCTATAGCGAGTTGTTGCTTGTGTCGAGGCGCAAACTTCGCGTGAAGGG 1499

QY 522 yAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa 542
DB 1500 AAACATACCAACAGCATGATCTATAGGAACACTGAGCAGTACCCTCATCAAGAACCGT 1559

QY 542 lProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrIle 562
DB 1560 TCCTGGTATCTCATCTTGGAGATTGATGCTCCCATCTACTTTGCTAATGCCAGTTACTT 1619

QY 562 uArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThrGlyGln 582
DB 1620 GCGTGAAGAATCATAGGTGGATTGATCAAGAGGAGAGAGAGTAAACAAATCAAGGAGA 1679

QY 582 uThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGln 602
DB 1680 GAGCAGCTTACAAATATATATATCTGATGTCAGTGTGTGTAATATTCGACACCAAGCGG 1739

QY 602 yIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLe 622
DB 1740 TATTAGCATGATGGTGAATTAAGAAAGTCATTCAGAGGAGCGGTTAAAGTTGCTATT 1799

QY 622 uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHi 641
DB 1800 GTCAAAATCAAAAGGAGGAGGTGTCGAAGAAATTAACACAGATCCAAATTCATCGGTGATCA 1859

QY 641 sLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAs 661

```

Db 1860 TTTGGGCAAGAGCTGGATGTTCTTAACGGTAGGAGAACAGTGGAGGCTTGTAGCTACAT 1919
Qy 661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrAsnVal 680
Db 1920 GCTTCACAGCTTTAAACCGAACCG--GCCTCAAAACGAGGCTTGGAAACAGTA 1974

RESULT 2

US-09-938-842A-1502
; Sequence 1502, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1502
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1502

Alignment Scores:

Pred. No.: 5,93e-263 Length: 1977
Score: 2595.50 Matches: 498
Percent Similarity: 87.73% Conservative: 81
Best Local Similarity: 75.45% Mismatches: 77
Query Match: 74.69% Indels: 5
DB: 11 Gaps: 3

US-10-762-049-18 (1-680) x US-09-938-842A-1502 (1-1977)

Qy 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
Db 1 ATGGGCAAGGAGGATACACATTCCTCAAGGAGCGGAGGAGCTTCACCGCCGCCACAC 60
Qy 44 ---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 61 ACGGTGGAGGCTCCACAACTCAACCGTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120
Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 121 ACTCTGTTCCAGACACCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 180
Qy 83 MetLeuGlyLeuGlnPhePheProLysPheGluTyrAlaProLysTyrThrPheGln 102
Db 181 GTGTAGGCGCTCAAAATCTCTCCCGATTTTCGAATGGGACACGCTCAATCTCAAG 240
Qy 103 PheLeuLysAlaAspLeuLeuAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 241 TTCTTCAATCAGATCTCATCGCGGAAATCACCATCGTAGTCTCGCCATCCCTCAGGC 300
Qy 123 IleSerTyrAlaLysLeuAlaAsnLeuProProLysLeuGlyLeuTyrSerPheIle 142
Db 301 ATCAGTTAGCCCAAACTTGCTAACTTGCCCAATTTCTGGCCCTTATTTCGAGTTTGA 360
Qy 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 361 CGCCCAATGGTATACCGGTGCTTACGGAGTTCAAGGGAGCTTAGCGGTGGAAACGGTGC 420
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182

Db 421 GTTGGTCTCTCTGTGACAGTCCGATGCTGAGCAAAAGAGTGTGATGCTGAGAAAGATCCT 480
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 481 AAGCTTTACCTTACCTTGTCTTACCGCCACTTTTTCGCCGGCGTTCGAAAGCTCT 540
Qy 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleG 222
Db 541 CTTGGAAATTTTCAGGTTAGGTTTCATAGTGGACTTTCTATCGCATGCAACGATAGTAG 600
Qy 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242
Db 601 ATTTCATGGAGGAGGAGCGACGCTGAGTCTGCAACAGCTTAAGGGTATTTCGGACT 660
Qy 242 uGluHisPheThrHisGlyAlaAspIleSerValMetArgSerValPheThrGlnTh 262
Db 661 TAAACATTTACAGACTTACCGATGTTATCTCTGTATCGCTTCGTTTTCCTCCAAAC 720
Qy 262 rHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
Db 721 TCACGAGTGGAGATGGAAAGTGGCTTCTTGGATGTGGTTTCTATTCTTCTCTCTC 780
Qy 282 rThrArgTyrPheSerLysLysArgProArgPhePheThrValSerAlaMetAlaProLe 302
Db 781 CACCAGATATTTGAGCATCAAGAAACCAAAATTTCTTTGGGTGGCGCGATGGCTCTTT 840
Qy 302 uThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVa 322
Db 841 GACCTCAGTGATTTCTTGGAGTCTCTTGGTGTCTTCTACTCACGCTGAGAGACATGGTGT 900
Qy 322 lGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342
Db 901 TCAAGTGATAGGGGACCTGAAGAAAGGTTGAATCCACTCTCGGTTCTGATCTCATCTT 960
Qy 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLe 362
Db 961 TACTTCCCTTACATGTCCACAGCTGTCAAAATGGCTCATCTCGGTCATCGGTCATCTCT 1020
Qy 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspG 382
Db 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTTTCGATGTTCAAGAACTACAAATAGACCGG 1079
Qy 382 yAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy 402
Db 1080 GAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGTTCTCTTCACATCTTGT 1139
Qy 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422
Db 1140 CCTCAACCCGACCATTTTCAAGGTCCGACGTGAACACTACAACGGGGTTCGAAGACCGC 1199
Qy 422 alaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLe 442
Db 1200 AATGTCCAAATAGTATGAGGCGATTTGCGGTTATGTTTACACTCTCTTCTCACACCGCT 1259
Qy 442 uPheHisTyrThrProLeuValValSerAlaIleIleValSerAlaMetLeuGlyLe 462
Db 1260 TTTTTCATACACACCATCTCTCTCTCCATCATCATATCGCAATGCTCGGACT 1319
Qy 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482
Db 1320 CATTGACTATCAAGCTGCGCATCTCTGGAAGTTCACAAAGTTCGACTTCTCTCTCTG 1379
Qy 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla 502
Db 1380 CATGAGCGCTTACGTTGGGTCGTATTTCGCGAGTGTAGAGATTGAGCTCGTTCGTCGT 1439
Qy 502 eValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuG 522
Db 1440 GCGGATATCTATAGCGAGGTTGTTCTGTTGTGTGAGGCCCAAAACTGCCGTGAGGG 1499
Qy 522 yAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa 542

```
Db      1500 AAACATACCAACAGCATGATCTATAGNAACACTGAGCAGTACCACCATCATCAAGAACCGT 1559
Qy      542 lProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrIle 562
Db      1560 TCCTGGTATTCTCATCTTGGAGATTGATGCTCCCATCTACTTTGCTAAATGCCAGTTACTT 1619
Qy      562 uArgGluArgIleThrArgTTrpIleAspGluIleGluArgIleLysAlaThrGlyG1 582
Db      1620 GCGTGAAGAAATCATAAAGTGATTGATGAAGAGGAAGAGAGATTAAACAATACAGGAGA 1679
Qy      582 uThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerG1 602
Db      1680 GAGCAGCTTACATATATTTACTCGATATGTTCAGCTGTGTGTAATATCGACACAGCGG 1739
Qy      602 yIleSerMetLeuGluGluValIleLysIleThrGluArgArgGluLeuGlnLeuValIle 622
Db      1740 TATTAGCATGATGGTGGAAATTAAGAAAGTCTATTGACAGGAGCGTTAAAGTTGGTATT 1799
Qy      622 uValAsnProValSerGluValMetIleLysIleAsnIleLysSerIleAspPhe--GlnAsnHi 641
Db      1800 GTCAATCCCAAAAGGAGAGGTCGTGAAGAAATTAACAGATCCAAATTCATCGGTGATCA 1859
Qy      641 sLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAs 661
Db      1860 TTTGGCAAGAGTGGATGTTCTTAACGTAGAGAGACAGTGGAGGCTTTGTAGCTACAT 1919
Qy      661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrAsnAsnVal 680
Db      1920 GCTTCACACGTTTAAACCGAACCG--GCCTCCCAAAAGAGCGCTTGGAAACAACGTA 1974

RESULT 3
US-10-437-963-38184
; Sequence 38184, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 38184
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRP4530_41845C.1
US-10-437-963-38184

Alignment Scores:
Pred. No.:      7,67e-235      Length:      2372
Score:          2330.50      Matches:    441
Percent Similarity: 83.84%      Conservative: 83
Best Local Similarity: 70.56%      Mismatches: 100
Query Match:      67.06%      Indels:    1
DB:              19          Gaps:      1

US-10-762-049-18 (1-680) x US-10-437-963-38184 (1-2372)

Qy      43 GlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db      149 AGGGTCCGATGCGCGCGAGAGCGGTTCTTGAGACGCTCGGGGGGAACATGAGGAG 208
Qy      63 ThrPhePheProAspAspProLeuArgGlnPhe--LysAsnLysProAlaSerLysLys 81
Db      1299 GCGATGTCGAACGTCATCATGTGCGGTGCGGTGATGATCAGCTGCTGTTCTTTCGCGCG 1348
```

```
Db      209 ACATTCTCTCGCGAGCACCGCTTCAGGGTGTGTGCGGCGGAGCGCGGGTGGCGGCGCGC 268
Qy      82 PheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPhe 101
Db      269 CGCGCGCGCGCGCTCCCGTACGTGTTCCTCGTTCATGAGAGTGGCGCGCGTCTGACACCCCTC 328
Qy      102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db      329 GGCACCTTCAGTCCGACCTCATCGCGCATCACCATGTCAGCCTGCCATCCCCAG 388
Qy      122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPhe 141
Db      389 GGCATCAGTCAGCAAGCTCGCAACCTCCTCCCGTCTCGGCTCTATTTCAGCTTC 448
Qy      142 IleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db      449 GTGCGCGCGCTGGTGTACGCGATGTGGGAGCTCGAGGGACCTGGCGGTGGGACCGGTG 508
Qy      162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db      509 GCGGTGGCGTCCGTGCTGATCGGGTCGATGCTGACGAGGAGGTGTCGCGCGCGGAGGAC 568
Qy      182 ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db      569 CCGCGCGCTGTACCTGCAGCTCGCGCTCACCGCCACCTTCTTTCGCGCGGTGTTCAGGCG 628
Qy      202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db      629 CTGCTGGGGGTCTTGAAGCTGGGGTTCATCTGTGACCTTCCTGTGCACGCCACCATCTCTC 688
Qy      222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db      689 GGGTTCATGGCGCGCGCCGCCACCTGCTGTGTGTCCTGACGAGCTCAAGGGCATGTTCGCG 748
Qy      242 LeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln 261
Db      749 CTCGACCACCTTCAACCCACCGCCACCGACCTCGTCTCCGTTCATGAGCTCGTCTTCTCCCAA 808
Qy      262 ThrHisGluTrpArgTTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeu 281
Db      809 ACCACCTCTGCGCATGGGAGAGCGTCTGTCATGGCTGCGGCTTCCTCTTCTTCCTCTCTC 868
Qy      282 SerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaPro 301
Db      869 ATCACCCCTCTTTCAGCAAGAGGAGGCAAGGTCTTCTGGGTATCTGCAGCTGCGCA 928
Qy      302 LeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGly 321
Db      929 TTGGCGTCTGTCTCATCATCGGAGCCTCTGCTGTACCTCACACATGCTGAAAAACCATGGC 988
Qy      322 ValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuVal 341
Db      989 ATTCAAGTGATTGGTTACTTGAAGAAGGGCCTGAACCCACCGCTCTGCACAAAGCTGAAC 1048
Qy      342 PheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSer 361
Db      1049 TTTCTGTCGCGGTATCATGATGCTGCCCTGAAGACCGGAGCATCATCTGGCGGTCAATTGCC 1108
Qy      362 LeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAsp 381
Db      1109 CTCGCTGAGGGGATTGCGGTAGGAAGGAGCTTTGCAATGTTCAAGAAATTACCACATCGAC 1168
Qy      382 GlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCys 401
Db      1169 GGAACAACAAGGAGATGATCGCGTTTGGAAACGATGTGGGAGTACTCACCCTCTTCTGTC 1228
Qy      402 TyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThr 421
Db      1229 TACCTCACACCGGTCGCTTCTCGAGGTGCGCGCTCACTACACGCCGCTGCAAGAG 1288
Qy      422 AlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrPro 441
Db      1289 GCGATGTCGAACGTCATCATGTGCGGTGCGGTGATGATCAGCTGCTGTTCTTTCGCGCG 1348
```

QY 442 LeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGly 461
Db 1349 CTGTTCCACTACACGCGCTGGTGGTCTGCGCGATCATCATGTCCGCGATGCTGGGG 1408
QY 462 LeuIleAspTyrGluAlaIleHisLeuPheLysValAspLysPheAspPheValVal 481
Db 1409 CTGATCGACTACCCGCGCGCTCCACCTGTGGCAGTGGCAAGGTGACTTCTCGCTC 1468
QY 482 CysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyValIleAla 501
Db 1469 TGCCTCGCGCCCTACCTCGCGCTGCTTCGCGAGGTGAGATCGGCTCGTGGCGCC 1528
QY 502 IleValIleSerValLeuArgValLeuLeuPheIleAlaAraProArgThrPheValLeu 521
Db 1529 GTCGGATCTCCATCTCCGCGTGTCTGCTGTCGTCGCGCGCGCGAGGAGCGGTGCTC 1588
QY 522 GlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHis 541
Db 1589 GGGAAACATCCCAACACGATGATCTACCGCGCGATGGACCACTACACCCCGCGCAGAGG 1648
QY 542 ValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyr 561
Db 1649 GTCCCGCGGTGCTGCTGCTCGCGTGCATCTGCGCTGCTGCTGCTGCTGCTGCTGCT 1708
QY 562 LeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThrGly 581
Db 1709 CTGCGTGAGAGGATCGCGGTGGATCGACGAGGAGGACCACTGCAAGGAGGCG 1768
QY 582 GluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSer 601
Db 1769 GAGATGGGCACTCAGTACGTCGCTCGACATGGGTGCGGTGCGAAGCATCGACACCGAGC 1828
QY 602 GlyIleSerMetLeuGluValLysIleThrGluArgArgGluLeuGlnLeuVal 621
Db 1829 GGGACAGGATGCTGGTGAATCAGGAAGACCTGGACAGGAGGCGCTTCAGATTGG 1888
QY 622 LeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHis 641
Db 1889 TTGGCAACCCGGGAGCGAGATCATGAGAGATTGGACAGCTCCCAAGGTGCTTGAGGCG 1948
QY 642 LeuGlyLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661
Db 1949 ATCGGCGATGAGTGGATCTTCCCAACGTTGGCGAGGCGGTGGCGAGTGGCTCGTG 2008
QY 662 LeuArgAlaSerLys 666
Db 2009 ATGCACTCGCAGAG 2023

RESULT 4
US-10-424-599-131116
; Sequence 131116, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131116
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89404C.1
US-10-424-599-131116
Alignment Scores:

Pred. No.: 7.31e-204 Length: 2008
Score: 2036.00 Matches: 403
Percent Similarity: 97.39% Conservative: 8
Best Local Similarity: 95.50% Mismatches: 9
Query Match: 58.59% Indels: 2
DB: 18 Gaps: 0
US-10-762-049-18 (1-680) x US-10-424-599-131116 (1-2008)
QY 261 GlnThrHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
Db 3 CAACGCATGAGTGGAGGTGTGAAGCGCTGTGTTGGGTGTGCTGTATTTTCTTCCTC 62
QY 281 LeuSerThrArgTyrPheSerLysArgProArgPhePhePhePhePhePheValSerAlaMetAla 300
Db 63 CTCGTCAACAAGTACTTTCAGCAAAACGACAGCAAGTTCCTCTGGGTGTGCAAAATGGCA 122
QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db 123 CCACGTGACGTCGCTTATATGGGAAGTCTGCTGCTATGTGTACACACACGCTGAGAAACAT 182
QY 321 GlyValGlu--ValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnL 340
Db 183 GGAGTTCAGGTGTATAGAGAAATTTGAAGAAGGGTTTGAATCCACCATCCTGACAAATC 242
QY 340 euValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleI 360
Db 243 TGGTATTTGTGACGCTTACATGACCAACAGCTGTCAAACAGGCAATGTGCTGGCATTA 302
QY 360 LeuSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnI 380
Db 303 TATCGTTCGCGAAGAAATAGCAGTGGGAAGAGCTTTTGCAATGTATATAAAATTTCAATA 362
QY 380 LeuAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrS 400
Db 363 TTGATGGCAACAAGAGATGATAGCTATTGGACCATGCAAGCTAGTGTGTTTCTTACCT 422
QY 400 erCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysL 420
Db 423 CTGTGCTACTCACAACTGACCACTTTTCGCGTTCAGCTGTGAACATATAACGCTGGATGCA 482
QY 420 ysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuI 440
Db 483 AGACAGCAGCATCCAACTCATATATGTCAATGTGCAATGTGCAATGTGTTATTTCTTAA 542
QY 440 hrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetL 460
Db 543 CACCAATGTTCCATTACACTCCCTCGGTGGTGTGCTATCAGCCATTAATTTGATCTGCAATGC 602
QY 460 euGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheV 480
Db 603 TTGGCTCATAGATTATGAGCAGCGCATCCATCTATTTAAGGTCGCAAAATTTGATTGTG 662
QY 480 alValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValI 500
Db 663 TCGTGTGATGAGTGCATACGTTGGCGTGTGCTTTGGCAGTGTTCGAAATTTGGCTTAGTCA 722
QY 500 leAlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheV 520
Db 723 TAGCTATTGGCGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782
QY 520 alLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaL 540
Db 783 TTTTGGGAAACATTCCAAAATTCGTGATATACAGAAATGTTGAGCAGCTATCCAAATGCAA 842
QY 540 ysHisValProGlyMetLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 560
Db 843 AACATGTTCTCGAATGCTAATTCAGAGATTGATGACCAATTTTACTTTTGTGCAATGCAA 902
QY 560 erTyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaT 580
Db 903 GCTATTTAAGAGAAAGGATCAAGATGGATTGATGAGGATGAGAGAAAGATCAAGAGCTA 962

```
QY 580 hrGlyGluThrSerLeuGlnTyrValIleAlaSerMetSerAlaValGlyAsnIleAspT 500
Db 963 CAGGGGAAACTAGTTGCGAGTATGTTATATTGATGATGAGTGCTGTTGCTTAACATTGATA 1022
QY 600 hrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnL 620
Db 1023 CAAAGTGGATATAGTATGCTTGAGAGGTGAGAGATTATAGAGAGAGAGAGATTACAA 1082
QY 620 euValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnA 640
Db 1083 TTGTTTGTGTCATCTGTAAGTGAGGTGATGAAGAACTGAACAAATCCAAGTTCCAAA 1142
QY 640 shHisLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 660
Db 1143 ATCATTTAGGGAAAAAATGGATCTATCTGACTGTTGAAGAGCGCTGGAGCATGCACT 1202
QY 660 heAsnLeuAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrAsnAsnV 680
Db 1203 TCATCTACGTGCAACCAACGACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1262
QY 680 al 680
Db 1263 TG 1264
```

RESULT 5

```
US-10-425-115-50601
; Sequence 50601, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 50601
; LENGTH: 2727
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MR74577_146143C.1
US-10-425-115-50601
```

Alignment Scores:

```
Pred. No.: 1.23e-184 Length: 2727
Score: 1856.00 Matches: 350
Percent Similarity: 73.20% Conservative: 117
Best Local Similarity: 54.86% Mismatches: 161
Query Match: 53.41% Indels: 10
Db: 20 Gaps: 2
```

US-10-762-049-18 (1-680) x US-10-425-115-50601 (1-2727)

```
QY 42 HisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLys 61
Db 303 TACAAAGTCGGGTCCCGGAGAGAACCTCTCGCGAGATCTCTGACGCGGTGAAG 362
QY 62 GluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLys 81
Db 363 GAGACGCTTCTCGCCACACCGCTGCGGCGAGTACAGGACCGCCAGGTCGCAAGAG 422
QY 82 PheMetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLysTyrThrPhe 101
Db 423 ATCTGGCTCGGCTGCGACACATCTTCCCGGTGCTCGAATGAGAGGCGCATTTACTCCCTC 482
QY 102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db 483 GGCAAGTTCAAGGGTGATCTCATCGCTGGGCTCACCATTCGCCGCTCTGCATACCCAG 542
```

```
QY 122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuLeuGlyLeuTyrSerSerPhe 141
Db 543 GACATCGGGTATTTAAAGCTTGCTAACTTGCAGCAGAGGGTGGGACTATACAGTAGCTTC 602
QY 142 IleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db 603 GTGCGCCTCTGATATACGAGTGTGAGGAGCTCCAGGGACATAGCATCGGTCCAGTG 662
QY 162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db 663 GCGGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
QY 182 ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db 723 CGCTCGAGTACAGCGCGCTAGCCTTCACAGCAACCTTTTTCGCAAGGGTCACTCAGGCA 782
QY 202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db 783 GCGCTCGAGTCTTCAGGCTAGGTTTATCATAGAGTTCTTGTCTCATGCTGCCATTGTC 842
QY 222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db 843 GGATTCATGGCGCGCTGCCATCACCATTCGCTTCAGCAGCTGAAAGGATTCCTCGGA 902
QY 242 LeuGluHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThrGln 261
Db 903 ATTGCAAACTTCACCAAGAAATCCGATATTGTATCTGTTATGAAATAGTTTGGGAAAT 962
QY 262 ThrHisGlu--TyrArgTyrGluSerAlaValLeuGlyCysValPheIlePheLeu 280
Db 963 GTTCACCATGGTGGATTTGGCAGACATACTGATAGGAGCGACCTTCTGCGCATTCCT 1022
QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheThrValSerAlaMetAla 300
Db 1023 CTGGTTGCCAAGTACATTTGGGAAAAGGATAAAAGCTCTTCTGGGTGTCGCAATCGCA 1082
QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaLysLysHis 320
Db 1083 CCTCTCACTCGGTGATCATATCCACATTTTGTGTATACATCACCTGCTGCGAGATGAAG 1142
QY 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
Db 1143 GCGGTGCAATTGTCAAGAACATAAGGAAGGCATCAACCCACCTTCAGCTAGTCTAATA 1202
QY 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
Db 1203 TACTTCACCGGCCCATACTTGGCGACAGGATTCAAAATTTGGGATAGTAGTGGATGATA 1262
QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
Db 1263 GGCCTTAACGGAAGCGATTGCAATTGGAAGAACATTTGCAAGCCCTCAAGAGATTACCGGATA 1322
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
Db 1323 GATGGGAACAAGAAATGATGGCTTAGGACCATGAACATTTGTTGGTTCAGTACTTCT 1382
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
Db 1383 TGCTACGTAGCCACAGGTTCTTCTCGCGGTGAGCAGTAAATTAATCATGGCTGGCTGCAAA 1442
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db 1443 ACAGCAGTGTCAACGTTGTTATGTCAAATGTGTAATGCTTACTTCTACTGATCACC 1502
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db 1503 CCATGTTCAAGTACACTCCAAATGCCATCTCTCTCGATCATCATATCAGCAGTGCTT 1562
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 1563 GGTTTAAATGACTATGAATCGGCTTACCTTATTCGAAAAGTTGACAAACTGGACTTCTTA 1622
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
```



```
Db 1623 GCATGCAAGGAGCATTTTGGAGTCATATTTTCATCGTGAGTATGCTTGCTCAIT 1682
Qy 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
Db 1683 GCGGTTGCAATATCTCTTGTAAATTTCTTCAAGTAACACGGCCAAAGACAGTTTAA 1742
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1743 CTTGGAAACCTTCCACGAACTATATACAGGAATGTAGAACAGTATCTGATGCTACC 1802
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db 1803 AAGGTTCCAGGGTGTGATTTAGAGTGGACTCAGCTATATATCTTCAAAACTCTAAC 1862
Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db 1863 TATGTTAAAGAGAGAAATCTCGAGTGGCTTAAGAGATGAGGAGGAGCAACACAGACCAG 1922
Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db 1923 AAGTTAAACAAACTGAGTTTCTAATTTGTGACTGTCTCTCTAATTGATATCGACACA 1982
Qy 601 SerGlyIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeu 620
Db 1983 AGTGGAAATCATGCTTTGGAGGAGTTGGCGAAAGCTCTTGAAAAACGCCAAAATTCAGCTG 2042
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db 2043 GTTCTTTACCAATCCCGGGCGCGGGTGTATCCAGAAGCTCGCTCAGCGAAATTCACGGAC 2102
Qy 641 HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPhe 660
Db 2103 ATGATTGGTGAAGACAAACATATTTCTCACGGTCGGCAGCCTGTG----- 2147

Qy 661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsn 678
Db 2148 -----AAGAAATTTGCTCTTAAGATGTCGATAATGCTGATGGAGT 2189

RESULT 6
US-10-425-115-182207
; Sequence 182207, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 182207
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97750C.1
US-10-425-115-182207

Alignment Scores:
Pred. No.: 1.5e-183 Length: 2651
Score: 1845.50 Matches: 355
Percent Similarity: 74.41% Conservative: 116
Best Local Similarity: 56.08% Mismatches: 157
Query Match: 53.11% Indels: 5
DB: 20 Gaps: 3

US-10-762-049-18 (1-680) x US-10-425-115-182207 (1-2651)

Qy 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
```

```
Db 518 GTGCACAAGGTGCGCGCGCGCGCGAGCACCGCGAGCAAGATGAAGGTGAGGGTG 577
Qy 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db 578 AAGGAGACCTTCTTCCCCGACGACCGTTCCCGGGGCTTCAAGGGCGGCGCGCGGAGCG 637
Qy 81 LysPheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThr 100
Db 638 CAGTGGCTCATGCGCGTCAAGTACCTTTCCTCCCATCTCGGACTGGGTGCGGAGTACTCC 697
Qy 101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
Db 698 TTGTCGCTCTTCAAGTCCGACCTCGTCGCGCGCTCACCATTCGCCAGCCTCGCCATTCCT 757
Qy 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140
Db 758 CAGGGCATTTAGCTACGCAAGCTGGCAAGCTTGGCTTCCCTATTCCTGCTGATTCGAGC 817
Qy 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db 818 TTGCTGCGCGCGCATGGGTACCGGTGCTGGGGAGCTCCCGTGACCTGCGGTGGCGCGCG 877
Qy 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
Db 878 GTGTCGATCTCGTCGCTGATCATGCGGTCCATGCTCGCGCAGCGCTGAGCCGCCACCGCG 937
Qy 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
Db 938 GAGCGCAGCCTGTTCTCGCAGCTGGCCTTTCACCTCCACCTGTTTCGCGGGCTGGTGCGAG 997
Qy 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
Db 998 GCTCTCGTGGGATCTCTCAGGCTCGGCTTCGTCATCGACTTCTGTCCAAGGCGAGCGCTG 1057
Qy 221 IleGlyPheMetGlyIleAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
Db 1058 GTGGGGTTCATGCGCGCGCGCCCATCATGCTGGCGCTGCAGCAGCTCAAGGCGCTGCTG 1117
Qy 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
Db 1118 GGCATGCTGCTCCATTCACCAACCGAGATGGGCATGCTCCAGTCATGGCCTTCGCTTCCAC 1177
Qy 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280
Db 1178 CACACGAGGAGTGTGCTGGGAGGATCCTCATGGGGCTGCTGCTTCTGCTTCTGCTG 1237
Qy 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
Db 1238 CTGTGCGCGAGCATGTGAGCATCAGATGGCCAAAGCTTTTCTGGGTTCGCGGTGCGCG 1297
Qy 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db 1298 CCCCTGGCATCGGTCAACATCTCGAGCTGTGTTTCTCTTCAAAAGCTCAGAAACCAT 1357
Qy 321 GlyValGluValIleGlyLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
Db 1358 GGCATCAGCATCATTTGGGCGAGCTCAAGTCGGCCTGAATCGCCCTCTGTGGGCAAGCTC 1417
Qy 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
Db 1418 CTGTTTGACACGCGGTATTTAGGCTCCATCAGAGCTGGCTTCTGCTCAGCGGAATCATC 1477
Qy 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
Db 1478 TCACTGACGGAAGGAATAGCGGTGGTAGAACATTTGCTCCTCACTCAAGGACATACCA 1537
Qy 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
Db 1538 GATGGAAACAAGGAGATGATGCGCATAGGCTTGAATGTTGTTGGGCTCTGCACATCA 1597
Qy 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
```



```
Qy 304 SerValIleLeuGlySerLeuValTyrPheThrHisAlaGluTyrHisGlyValGlu 323
Db 966 TGTGTATCTCTACCCCTTGTGTTTTCGAATTAAGGCTCAAAATCATGGCATCAGT 1025

Qy 324 ValIleGluLeuLysGlyLeuAsnProSerLeuThrAsnLeuValPheVal 343
Db 1026 GCGATTGGAATTTGCAACAGGAATAATCCTCCATCATGGAATATGTTGCTCTTTTCAT 1085

Qy 344 SerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAla 363
Db 1086 GGAAGTCACCTAGCCCTAGTTATGAAACAGGGCTTATCACCGGATTTTGTCCCTTAACA 1145

Qy 364 -GluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAs 383
Db 1146 GGAAGGTATTGTCAGTAGGAAGACATTTGCAGCTCTCAAAACACTACAAAGTGATGGAAA 1205

Qy 383 nLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLe 403
Db 1206 TAAGGAATGATGGCAATGGGTTTATGAAATGTTTGGTCCCTTCACCTTCCTGCTATGT 1265

Qy 403 uThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAl 423
Db 1266 TACAACAGGTGCTTTCTCTCGGTGAGCAGTTTAAACAATGCGAGCGGCGCAAAACAGCTGT 1325

Qy 423 aSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPh 443
Db 1326 GTCAAATGATGATGTCGTGACAGTCATGTTGACACTCTTTTCTCATGCCATGTT 1385

Qy 443 eHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuI 463
Db 1386 TCAATACACGCCATAATGCTGCTGGCGCAATCATAGTCACAGCAGTAATTCGCTCAT 1445

Qy 463 eAspTyrGluAlaAlaIleHisLeuPheTyrValAspLysPheAspPheValValCysMe 483
Db 1446 CGATCTCCCGCTGCTGTGAACATTTGGAAGATCGACAAATTCGATTTTGTGTGATAT 1505

Qy 483 tSerAlaTyrIleGlyValValPheGlySerValGlyLeuValIleAlaIleVala 503
Db 1506 GACTGCTTTCTTAGTGTTCTTTTCTCTCTGTCGAAGGCGCTGCTCTGCTGTGG 1565

Qy 503 lIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAs 523
Db 1566 ATTATCAACTTTGAAGATACTCTTGCAAAATTACGAGACCGCAAAACAGTGATGTTGGGAA 1625

Qy 523 nIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValPr 543
Db 1626 GATACCGAGGAACAGACATATATAGAAATCTTGATCAATACAGGAAGCTGTGAGAAATACC 1685

Qy 543 oGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuAr 563
Db 1686 TGAATTTCTTATTTTAAGCATTCAGGCTCCCATCAATTTTGCCAAATCATCATATCTCAA 1745

Qy 563 gGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThr 583
Db 1746 CGAGAGAACGTTAAGATGGATTGAAGAAGAAGAACAAATATAAGAA-----CAATT 1799

Qy 583 rSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyI 603
Db 1800 AACCCCTTCGATTCCTTAGTATTGAAATGTGCAGCTGTGAGTGTGCTTTGACACAGTGGAA 1859

Qy 603 eSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuVa 623
Db 1860 CTCACCTTTCAAGAGATTGAAACCAACACTGGAAGAAAGAGGTGTTGAGCTGTGTGTGT 1919

Qy 623 lAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn----HisLe 642
Db 1920 CAATCTCTTGTGTGAGTTCATAGAAAAGCTTTAAAAAAGCAGATGAAGCTAATGATTTTCAT 1979

Qy 642 uGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLe 662
Db 1980 AAGAGCAGATAACCTTTTCTTGACAGTTGGAGAGGCTGTAGCT-----TCACT 2027
```

```
Qy 662 uArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 2028 TTCTTCAGCAATGAAAGGCCAATCATCAACCATTCAGAAGGGGCGCACACAAT 2082

RESULT 8
US-10-437-963-95853
; Sequence 95853, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 95853
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94005C.1
US-10-437-963-95853
```

```
Alignment Scores:
Pred. No.: 3,7e-180 Length: 2001
Score: 1811.50 Matches: 340
Percent Similarity: 73.10% Conservative: 122
Best Local Similarity: 53.80% Mismatches: 167
Query Match: 52.13% Indels: 3
DB: 19 Gaps: 3

US-10-762-049-18 (1-680) x US-10-437-963-95853 (1-2001)
```

```
Qy 43 GinValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuYsGlu 62
Db 106 GAGGTGAACCTGTCCGGGGCGCGCGTTCGCGGAGAGCTGTGGTCGACCTCGCGGAG 165

Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProLaserLysLysPhe 82
Db 166 ACGTTCCTTCCCGCAGCACCCGTTCCGCGGGGTTCCGCGCGCTGCCCGCCAGCGCGCG 225

Qy 83 MetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLysTyrThrPheGln 102
Db 226 TGTGTCGGGTCAAGTACTTCGTCGCGGCTGGAGTGGTCCGGGCTACGGCCTCGAC 285

Qy 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 286 AAGTTCACAGTTCGACTCTCTCCCGGCATCACCATCGCCAGCTCGCCATCCCGCAGGCG 345

Qy 123 lIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIle 142
Db 346 ATCAGCTAGCGCGCTCGCCAACTCCCTCCAACTCATCGGCTCTTATTCGAGCTTCGTG 405

Qy 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 406 CCGCGCTGATGATGACGGGTTCGAGAGCTCAACACCTGGCGGTGGGAGCGGTGGCG 465

Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
Db 466 GCGCGCTGCTGCTGTCGCTCATCATCGAGACGAGGTCGCGCGGACGAGACCCG 525

Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 526 CAGCTGTACTCAGCTCTTCTTACACCGCGCTTCTTTCACCGCGCTCTTCTCCAGACCGCG 585
```



```

Db      613  GACTGGGGCGGCACCTTCCGCAAGTTTCAGGGCGACCTCGTCTCCGGCCCTCAC 672
Qy      114  IleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuPro 133
Db      673  ATTGCCAGGCTTCGCATACCCAGGACATCGGTATGCGAAGCTTCTGCGCCATTATACA 732
Qy      134  IleLeuGlyLeuTyrSerSerPheIleProProLeuIleTyrAlaMetMetGlySer 153
Db      733  AATTATGGAGCTACAGCAGCTTCGTACCGCTTTTGATATACGCGATGATGGAAAGTTCC 792
Qy      154  ArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMetLeu 173
Db      793  AGGATATAGCATTGCTCCAGTGGCGCTGTTTCGCTGCTGCTGGCAGCTCTCTCCAG 852
Qy      174  AsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAlaPheThrAla 193
Db      853  AATGAGTTTGTATCCCAAGAGAAATCAGGAAGAGTACACCCGGCTGGCCCTTCAGTCA 912
Qy      194  LeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGlyLeuIleVal 213
Db      913  TTCTTCGCTGGGTCAACCCAGGAGTGTCTGGATTCTCTCAGCTAGGTTTATCATAG 972
Qy      214  PheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValValCys 233
Db      973  TTCTTCGCTGCTGCCATCGTTCGATTCATGGCGGAGCGCCATCACCATTTGCCCTT 1032
Qy      234  GlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAspIleLeu 253
Db      1033  CAGCAGCTTAAAGGCTTCCTTGGAAATGCAAACTTCACCAAGAGACTGACATCATCT 1092
Qy      254  ValMetArgSerValPheThrGlnThrHisGlu---TyrArgTyrPheLeuValLeu 272
Db      1093  GTCATGAATCATGTGGGGAATGTTTACCATGGTGGAACTGGCAGACGATATGATC 1152
Qy      273  GlyCysValPheIlePheLeuSerThrArgTyrPheSerIlyLysArgProArg 292
Db      1153  GGAGCATCTTCTCGCATTCCTCTCTGTTGCAAAATACATTGCAAAAGAAATAAGAA 1212
Qy      293  PhePheTyrValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuVal 312
Db      1213  CTCCTTTGGGTGCTGCATATGCACCATCTCCTCGGTGATCATATCCACCCCTTTT 1272
Qy      313  TyrPheThrHisAlaGlyLysHisGlyValGluValIleGlyCyluLeuLysLysGly 332
Db      1273  TATATCATCGACGACGACGATGTTGCTGATTTGTCAGTACATACAAAGAAAGGCATC 1332
Qy      333  AsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaVal 352
Db      1333  AATCCACCTTCAGCTAGTCTCATATATCTTCAGTGGCCCAAACTGTATGAAGATTCA 1392
Qy      353  ThrGlyLeuValValGlyIleLeuSerLeuAlaGluGlyIleAlaValGlyArgSer 372
Db      1393  ATAGGAGTATAGCTGGCATGATAGGCTTAACAGAAAGCAATTCGCAATAGAGAAAT 1452
Qy      373  AlaMetTyrIleAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThr 392
Db      1453  GCTGGCTGAGGACTACAGATTGACGGAACAGAAATGTTGCTCTAGAAACATG 1512
Qy      393  AsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSer 412
Db      1513  AACATTGTTGTTCAATGACTCTTCTAGTACGTTACAGGTTCTTTCTCGCGGTACGA 1572
Qy      413  ValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleLeuMetSerLeuAla 432
Db      1573  GTCAATTATACATGCTGGCTGCCAAACAGCGGTATCCAAATATTTGTTATGTCAAT 1632
Qy      433  MetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeu 452
Db      1633  CTGCTCACACTGGAGTGAATTTCTTCCATTGTTCAAGTATACGCCAATGCCATCTT 1692
Qy      453  AlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeu 472

```

```

Db      1693  TCCATCATCATATCAGCAGCTGCTGGCTTAGTCTGACATACACAGCATACCTTATCT 1752
Qy      473  LysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPhe 492
Db      1753  AAGTTGATTAACATTGACTTCTGCGATGCTTAGGAGCATCTTTGGCGTCATATTTCA 1812
Qy      493  SerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuPhe 512
Db      1813  TCAGTGGAGTATGGTCTGCTAAATTGGCTTCTCACTCGCTAAATAATCTTCTACAA 1872
Qy      513  IleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArg 532
Db      1873  GTTACACAGCCACAGAACAGTTTGTCTGTAACTTCCAGAACACATATATAGGAAT 1932
Qy      533  ValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAsp 552
Db      1933  ATCGACCATGATCCAGAGCTACCTTGTGTACAGGAGTGGTAATAGTTAGGTTGACT 1992
Qy      553  ProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTyrPheAsp 572
Db      1993  GCTATCTACTTTCACAAACTCGAACTATGTTAAAGACAGAACTCTGAGGTGGCTGAG 2052
Qy      573  GluGluGluArgIleLysAlaThrGlyLysSerLeuGlnTyrValIleIleAspMet 592
Db      2053  GAGGAAGAACGTCAGCAGAAACAGAGCTACAGAAACTGAGTTTCTGATTTGTGAGCT 2112
Qy      593  SerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLysLys 612
Db      2113  TCTCGGTAAATTGACATTGATACAGTGGATCCATGCTTGGAGGACTATTTAGAGCA 2172
Qy      613  ThrGluArgArgGluLeuGluValLeuValAsnProValSerGluValMetLysLys 632
Db      2173  CTTGAAAGCGCAAAATCCAGCTGATTTAGCAATCCGGGCGCCAGCTGTGATCTCTGA 2232
Qy      633  LeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTyrPheTyrLeuThrVal 652
Db      2233  CTCCGATCAGCGAAATTCACGGATCTCATCGTGAAGACAAAGATATCTCTGACTGT 2292
Qy      653  GluAlaVal 655
Db      2293  GACGAGTG 2301

RESULT 10
US-10-425-114-15719
; Sequence 15719, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15719
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-059-F10_FLI
US-10-425-114-15719

Alignment Scores:
Pred. No.: 9, 9e-175 Length: 2656
Score: 1762.00 Matches: 334
Percent Similarity: 73.23% Conservative: 120
Best Local Similarity: 53.87% Mismatches: 162
Query Match: 50.71% Indels: 4

```

```
DB: 18 Gaps: 2
US-10-762-049-18 (1-680) x US-10-425-114-15719 (1-2656)
QY 41 ValHisGlnValGluValProProGlnProPhePheLeuSerLeuLysTyrSerLeu 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 CTGCACAAGGTGTCGTCGCGGCGCGGACGACGCGGCGAAGCGCTGCGGCGCGCTG 358
QY 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 GCCGAGGTGTTCTCCGCGACGACCGCTCGCACCACTTCAGAGACGAGTCGTCGGCGGG 418
QY 81 LysPheMetLeuGluGlnPhePhePhePhePhePhePheGluTrpAlaProLysTyrThr 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 CGCCTCGTCTGGCTGACACTTCTTCCCATCTTCCAGTGGGGGTTCGCCCTACAGC 478
QY 101 PheGlnPheLeuLysAlaAspLeuAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 479 CGCGGCTCTCGGCTCCGACCTCGTCGCGGCGCTCACCATTCGCCAGCTCGCCATCCG 538
QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSer 140
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 539 CAGGGAATCAGCTACGCCAAGCTCGCAACCTCGCGCCCAATCGTTGGCCTATATTCCAG 598
QY 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 599 TTCGTGCCCGCGCTCATCTACGCGCTGCTGGGAGCTCGCGGAGCTGCGCGTGGGGCG 658
QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 659 GTGTCCATCGCTCGCTGGTGATGGGTTCATGCTCCGGGACGCGTGTCCGCGACGAG 718
QY 181 AspProLysLeuTyrIleuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 719 CAGCGCTCTCTACCTGAGCTGGCGCTTCACGCGCACCTTCTTCGCGCGGCTCTTCCAG 778
QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 779 GGTGCTCCCTGGGATTCCTCAGGCTGGGCTTCATCGTGGAGCTTCTGTCAAAGCGACGCTG 838
QY 221 IleGlyPheMetGlyValAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 839 ACGGGCTTCATGGCGGCGCGCCGCTCATCGTGTGCTGCAGCAGCTCAAGGGCGCTGCTC 898
QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThr 260
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 899 GGCATCTCCCACTTCACTCCACATCGGATTCCTCGAGCTCATCGGCTCCGCTCGTCAAC 958
QY 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 959 CGCCACGACGAGTGAAGTGGCAGAGATCGTCATGGGCTCGCGCTTCTCGCCATCTCTC 1018
QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1019 CTCCTCAGCGCCAAATCAGCCAGCAACCCAAAGCTTTCTGGGTATCAGCAGGTGCT 1078
QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1079 CCCCTGGCGTGGTATCATCTCCACCATCTCTCTCTCATCTGG-----AAATCCCCC 1132
QY 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1133 AGCATCAGGTATTGGCATCTCCCCAGGGAGTGAAACCTCTTTCGGCGGAACATGCTC 1192
QY 341 ValPheValSerProTyrThrThrAlaValLysThrGlyIleValValGlyIleIle 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1193 AGCTTCAGCGGCTCTATGTGGCGTGCAGATCAAAACCGGATCATGACAGCATCTCTG 1252
QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrIleAsnTyrAsnIle 380
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1253 TCCTTAACAGAAGGGATCGACGTGGGCGAGGACCTTCGCGTCCATCAACAACCTACCAGGTG 1312
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
```

```
Db 1313 GACGGGAACAAGAGATGATCGCATCGGGCTGATGAACATGGCGGGCTCTCGCGCTCC 1372
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1373 TGCTACGTGACGACGCGGGTCTCTTCCCGGTGCGGGTGAACATACGCGCGGGTGCAGG 1432
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuPheLeuThr 440
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1433 ACGCGCTGTCTCAACGTCGTATGCGGCGCGGTGTGGTGGAGCGCTGCTTCTCTCATG 1492
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1493 CGCTGTTCCTACTACACCCGAAAGTATCTCGGGGGATCATCATCACCGCGGTGGTG 1552
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1553 GGGCTGTGGAGCGTGGCGGCGCGCCAGGCTGTGGAAGGTGGACNAGCTGGACTTCCTG 1612
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1613 CGCTGCGTGGCGGCGTTCCTCGGCGTGTCTGTGTCTCGACAGCGGCGCTGGCGCTC 1672
QY 501 AlaIleValIleSerValLeuLeuLeuPheIleAlaArgProArgThrPheVal 520
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1673 GCGGTGCGCATCTCGCTTCAAGGTCTGTGTGAGGTACCCGCCCAACGTCGTGGTG 1732
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1733 GAGGCGCTCGTCCGCGGACGACAGAGCTACCGCAGCGTGGCGCAGTACCGCGAGCGCTC 1792
QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1793 CGCTGCGCGGCTTCTCTGCTGCGGCTCGAGTCCGCGCTACTTTCGCAACTCCATG 1852
QY 561 TyrLeuArgAlaGlyIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1853 TACTGTGGAGCGGTATCGGTACTCTCGCAGCAGGAGGAGCGCGCTCAAGTCC 1912
QY 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1913 AACCAACCCTCCATCCGATGCGTCTCTCGACATGGCGCGCTCGCGCGATCGACAG 1972
QY 601 SerGlyIleSerMetLeuGluValLysLysIleThrGluArgGluLeuGlnLeu 620
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1973 AGCGGTCTAGACGCGCTGTCCGAGCTCAAGAAAGTCTCGSACAAAGAAACATCGAGCTG 2032
QY 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2033 GTGCTTGCACACCGGTGGGTTCGTGGCGGAGAGATGTTCAACTCGCGCGTGGCGAG 2092
QY 641 HisLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaVal-----GlyAlaCys 658
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2093 AGCTTCGGGTGCGGCGCGCTCTTCTTACGCGTAGCGGAGGCGCTCGCGCGGGCGGTGC 2152

RESULT 11
US-10-425-115-75763
; Sequence 75763, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 75763
; LENGTH: 2713
; TYPE: DNA
```

US-10-437-963-73410/c
; Sequence 73410, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.


```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 73410
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73694C.1
US-10-437-963-73410
```

```
Alignment Scores:
Pred. No.:      6,85e-172      Length:      2429
Score:          1734.50        Matches:      333
Percent Similarity: 70.93%      Conservative: 116
Best Local Similarity: 52.61%    Mismatches:   163
Query Match:      49.91%       Indels:       21
DB:               19          Gaps:         3
```

US-10-762-049-18 (1-680) x US-10-437-963-73410 (1-2429)

```
QY 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2261 CTGCACAAAGGTCTCGGTGCGGAGAGAGGTTCGAGCGCGAAGCGCTGAGCGCGGTG 2202

QY 61 LysGluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2201 GCGGAGGTGTTCTTCCCGACGACCGCTCCACAGTTCAAGAACCAAGTCGTCCGCGCG 2142

QY 81 LysPheMetLeuGlyLeuGlnPhePhePheProLlePheGluTrpAlaProLysTyrThr 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2141 CGCGTGGTCTCGCGTCTGAGTACTTCTTCCCATCTTCCATTTGGGGTCCGACTACAGC 2082

QY 101 PheGlnPheLysAlaAspLeuLleAlaGlyLleThrLleAlaSerLeuAlaLlePro 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2081 CTCGCGCTCTCCGCTCCGAC----- 2061

QY 121 GlnGlyLeuSerTyrAlaLysLeuAlaAsnLeuProProLleLeuGlyLeuTyrSerSer 140
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2060 ---GGNATCAGTACGCCAAGCTCGCAACCTTGCCTCCCATCATTTGGACTATATTCGAGC 2004

QY 141 PheIleProProLeuLleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2003 TTCTGTCGCGCGTGTATCTACTCTGTGTGGGTAGCTCGCGGATCTGGCTGTAGGCGCG 1944

QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1943 GTGTGATACGTCGCTGTGTATGGGTTCGATGCTCCGCGAGCGGTGTGCGCGACCCAG 1884

QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1883 GAGCCCATCTTACCTCCAGCTCGCTTCACTCCACCTTCTTCCGCGCGCTTCTCCAA 1824

QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuLleValAspPheLeuSerHisAlaThrLle 220
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1823 GCTTCCCTCGGCTTCTCCGCGTGGGTTCATCGTGGATTTCTTTCGAAGGCGACGCTG 1764

QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLysSerLleLeu 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1763 ACGGGATTATGGGCGCGCGCGATCATCTGTCCTCGAGCGAGCTCAAGGGATTGCTC 1704

QY 241 GlyLeuGluHisPheThrHisGlyAlaAspLleIleSerValMetArgSerValPheThr 260
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1703 GGGATCATCAATTTCAGTCGCGAGATGGGGTTCGTCCTCAGGTGATGCACTCCGCTTCAG 1644
```

```
QY 261 GlnThrHisGluTrpArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1643 CACCACGACGAGTGGCGTGGCAGCACCATCTCACTAGGGCGTCGCTTCTCGCGCTCTC 1584

QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1583 CTCACACACGCCACATCAGCGCCAGNACCAAGCTTTTCTGGGTGTGACGAGTCT 1524

QY 301 ProLeuThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHis 320
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1523 CCACCTGACATCAGTGTATCTTACCATCATCTCGTTTCGTTCAGCAAGCT-----CAT 1470

QY 321 GlyValGluValIleGlyLeuLysGlyLeuAsnProProSerLeuThrAsnLeu 340
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1469 GGCATCAGTGTATTGGCGATCTCCCAAGGATTTGAACCTCTTTCAGCAATATGCTG 1410

QY 341 ValPheValSerProTyrMetThrAlaValLysThrGlyLleValValGlyLleIle 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1409 ACCTTCAGTGGCTCTAGTAGGACTGGCTTAAACACAGGATTTATGACTGGCATCTA 1350

QY 361 SerLeuAlaGluGlyLleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1349 TCTCTCACTGAAGGATAGCAGTAGGACAGACATTTGCATCCATCAACAACCTACCAAGTT 1290

QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1289 GATGGAAACAAGGAGATGATGCCATTGGTGTCTATGAACATGGCTGCTCATGTGCCCTCC 1230

QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1229 TGCTATGTACACACAGGATCTTCTCAGAGTTCGGCGTGAACTACAGCGCGGGTGCAAG 1170

QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1169 ACGCGGTGTTCGAACATCGTATGCGGTCCGCGGTGCTGTGTGACGCTCTGTTCCTGATG 1110

QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaLleIleValSerAlaMetLeu 460
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1109 CGCTGTTCCTCACTACACCCCAAGTGATCTCTCGCGGATCATCATCACCCTCGGTATC 1050

QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1049 GGCCTCATCGAGTCGCGGCGCGCCAGCTGTGGAAGGTTCGACAAGCTCGACTTCCTC 990

QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 989 GCCTGATGGCGCGCTTCTCGCGTCTCTCGTCTCGTCCAGATGGCGCTCGCCATC 930

QY 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaAtgProArgThrPheVal 520
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 929 GCGGTTCGCATCTCCCTTCAAGATCTCTCTCAGGTCCACCCGCCCAACATGTGTGTC 870

QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 869 AAGGCGCTGTCGCCGCGCAGCTACCGAGCATGGCGCATCAGGAGGCCATG 810

QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 809 CGTGTGCGCTCTCTCGTCTCGGTCTCGAGTCCGCGCATCTACTTCGCCAATCTCCATG 750

QY 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 749 TACCTCGCGAGAGGATCATGAGTTTCTCCGGAGGAGAGACGAGCGCGGCCCAAGTGC 690

QY 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 689 AACAGTCCCTGTGATGATCATCTCTCGACATGAGTGTGTGTCAGCGATTCGACAGC 630

QY 601 SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluGlnLeu 620
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 629 AGTGGCTCGATTCGACAGCTGAAGAAGGTGCTGGAAAAAAGAAACATCGAGCTT 570
```


QY 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db 569 GTGCTGCCAACCCGGTGGATCGGTGACCGAGAGGCTGTACAACCTCGGTGGTCGCAAG 510
QY 641 HisLeuGlyLysLysTrpIleThrLeuThrValGluGluAlaValGlyAlaCysAsnPhe 660
Db 509 ACCTTCGGCTTCAGACCGCGGTGTTCTTCAGCGTCCGAGGCGGTCCGCGCGCG 456
QY 661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGlu 673
Db 455 -----CCGCACAAACACAGCCCTGCACAGAGCAG 426

RESULT 13

US-10-437-963-77012
; Sequence 77012, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77012
; LENGTH: 4390
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76952C.1
US-10-437-963-77012

Alignment Scores:

Pred. No.: 1,296-167 Length: 4390
Score: 1698.00 Matches: 328
Percent Similarity: 72.61% Conservative: 120
Best Local Similarity: 53.16% Mismatches: 164
Query Match: 48.86% Indels: 6
DB: 19 Gaps: 2

US-10-762-049-18 (1-680) x US-10-437-963-77012 (1-4390)

QY 42 HisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLys 61
Db 503 TACAATGTGGCGCGCGCGAGAGAGAACCTCTCGCGGAGTTCGCGCGCAGGTGAAG 562
QY 62 GluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProLysLys 81
Db 563 GAGACGTTCTCCGACGAGCGATGCGCGGTACAGGACCAGCGGAGGTCCAGGAAG 622
QY 82 PheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPhe 101
Db 623 CTATGGCTCGCCTTGCAGCAGCGTTCCTCGCGGTTCGAATGGGCGAGACAATACACCTC 682
QY 102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db 683 GCCAAGTTCAGGGCGACCTCATTCGCGGCTCACCTTGCACGCTCGTCATACCTCAG 742
QY 122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerPhe 141
Db 743 GACATCGGCTACGGAAGCTTGTAACTGCCACAGAGATTGGGCTGCACAGTAGCTTC 802
QY 142 IleProProLeuIleTyrAlaMetMetGlySerArgAspLeuAlaValGlyThrVal 161
Db 803 GTCCCGCCCATGATATACGCTCTGTGGGCGACCTCAAGGGAGCTAGCAATGGGTCCAGTG 862

QY 162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db 863 GCGGTCTCATCTACTGCTGCTGCTCTCTCCAGGAGGAGTACTCAAAAGAAGAAC 922
QY 182 ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db 923 CCGCTAGATTACAGACGGCTCGCCTTCACAGGAGCTCTCTTTCAGGAGGTACACAGGCG 982
QY 202 AlaLeuGlyLeuPheArgLeuLeuValAspPheLeuSerHisAlaThrIleIle 221
Db 983 GCGCTGGGTTCGCGAGGTTCATCATAGGCTTCCTGCTCATGCTGCCATCATC 1042
QY 222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db 1043 GGATTCATGGCGCGCGCGCCCATCACCATTGCTCTTCAGCAGCTTAAAGGCTTCTTGA 1102
QY 242 LeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln 261
Db 1103 ATTGCAAACTTCACCAAGAAGACTGACATCATCTCGGTCAATGAATCAGTCTGGGAAAT 1162
QY 262 ThrHis-----GluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePh 279
Db 1163 GTTCACCATGGGCAATGGAAGTGGCAGACATATTGATCGGAGCATCATTTTGGCAT 1222
QY 279 eLeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMe 299
Db 1223 CTTCTCTGGT-TGCCAA-----GGCAAGAAGAACAAAGAGCTCTTCTGGGTCCAGCAAT 1275
QY 299 talProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLys 319
Db 1276 TGCACCACTCATTTTCGGTGATCATTTCACTTTTGTTCGTCTACATCACTCTGGTGCACAA 1335
QY 319 eHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAs 339
Db 1336 ACAGGCGTGGCAATCGTAAACAGTCAAGAAAGCATCAATCACCTTTCAGTAGCT 1395
QY 339 nLeuValPheValSerProTyrMetThrAlaValLysThrGlyIleValValGlyIle 359
Db 1396 GATATTTTCTACTGGCCCATCTTCTCAAGGATTCAAAATTGGAGTAGTAGTGAAT 1455
QY 359 eIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAs 379
Db 1456 GATAAGCTTACGGAAGCGATTGCGAGTTGGAAGAACATTTGCTGGAATGAACGATTACCA 1515
QY 379 nIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheTh 399
Db 1516 GATAGATGGGAACAAAGAAATGTTGGCTCTAGAACCATGAATGTTGGTTCATATGAC 1575
QY 399 rSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCy 419
Db 1576 GTCTTGTATATAGCCACAGGTGGTTCGACGATCAGCAGTCAATTCGATGGTGGAGG 1635
QY 419 sLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLe 439
Db 1636 TAAACACCAATATGTCATATTTGTTATGTTATGTTACCTAGTAGTATTGCTTGCACTCT 1695
QY 439 rThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMe 459
Db 1696 CACTCCATTTTCAAGTACAGCCCAATGCCACCATTTCTTCCATCATCATATCAGCAT 1755
QY 459 tLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPh 479
Db 1756 GCTTGGCTATTGATTTGATCTTGAATCAGCTACCTTATCTGGAAGTTGATAAGTTGAC 1815
QY 479 eValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVa 499
Db 1816 CATGCGGTCTTGGGCGCATCTCTTGGAGTAATAATTTTCATCTGTGGAGTATGGCTTGT 1875
QY 499 tIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPh 519
Db 1876 CATTCGGGTGTAATATCACTAATCAAAAGTTCTGCTCCATGATGTAACACGCCCAAGGAC 1935
QY 519 eValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAl 539


```
QY      580 ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp 599
Db      1696 GTCCCGCGGAGGACCTGCTACGTCGTCTCGACATCGCGCGCTGACGGCGATCGAC 1755
QY      600 ThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGln 619
Db      1756 AGCCCCGGGATCGAGATGCTCGGGAGGTGCACGGCGAGCTGGAGAGGAAGGGGATGAAG 1815
QY      620 LeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln 639
Db      1816 ATGGCGGTACCAACCCGAGGATGGCGTGGCGGAGAGCTGTGCTGTCCGGGCTCGCG 1875
QY      640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db      1876 GAGCTGTCGGCGAGAGCTGGATGTTCTCTCCAATGGCGACGCTTGGCCGCGTGC CGG 1935
QY      660 PheAsnLeuArgAlaSerLys 666
Db      1936 TACACGCTCCAGGGCTCCAAG 1956
```

Search completed: September 12, 2005, 21:40:45
Job time : 1049 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 31, 2005, 09:44:04 ; Search time 8026 Seconds
(without alignments)
4105.351 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGNNV 680

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10762049/runat_30082005_121245_5972/app_query.fasta_1.839
-DB-GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10762049 @CGN 1 1 4545 @runat_30082005_121245_5972 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_sm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	3475	100.0	2449	6 AR477464 Sequence
2	2642.5	76.0	2045	8 D89631 Arabidopsis
3	2641.5	76.0	2266	8 AY099873 Arabidopsis
4	2597.5	74.7	2193	8 BA581745 Brassica

	5	2595.5	74.7	1977	6	AX506807	Sequence
	6	2595.5 <td>74.7</td> <td>1977</td> <td>6</td> <td>AX651748<td>Sequence</td></td>	74.7	1977	6	AX651748 <td>Sequence</td>	Sequence
	7	2330.5 <td>67.1</td> <td>1974</td> <td>6</td> <td>AX653455<td>Sequence</td></td>	67.1	1974	6	AX653455 <td>Sequence</td>	Sequence
	8	2330.5 <td>66.1</td> <td>2327</td> <td>8</td> <td>AK066932</td> <td>Oryza sat</td>	66.1	2327	8	AK066932	Oryza sat
	9	2314	66.6	2037	8	AB004060	Arabidops
	10	2274.5 <td>65.5<td>2219</td><td>8</td><td>AJ601439</td><td>Brassica</td></td>	65.5 <td>2219</td> <td>8</td> <td>AJ601439</td> <td>Brassica</td>	2219	8	AJ601439	Brassica
	11	2164	62.3	1981	6	AR477457	Sequence
c	12	2026.5 <td>58.3</td> <td>123209</td> <td>8</td> <td>AT4F415</td> <td>Arabidops</td>	58.3	123209	8	AT4F415	Arabidops
c	13	2020.5 <td>58.1</td> <td>60000</td> <td>8</td> <td>AF049236</td> <td>Arabidops</td>	58.1	60000	8	AF049236	Arabidops
	14	2012.5 <td>57.9</td> <td>7012</td> <td>8</td> <td>AB012048</td> <td>Arabidops</td>	57.9	7012	8	AB012048	Arabidops
	15	1901	54.7	2067	6	AR477463	Sequence
	16	1898.5 <td>54.6</td> <td>1896</td> <td>6</td> <td>AX651700</td> <td>Sequence</td>	54.6	1896	6	AX651700	Sequence
	17	1894.5 <td>54.5</td> <td>1917</td> <td>8</td> <td>AB023423</td> <td>Arabidops</td>	54.5	1917	8	AB023423	Arabidops
	18	1892.5 <td>54.5</td> <td>2470</td> <td>8</td> <td>AF347614</td> <td>Arabidops</td>	54.5	2470	8	AF347614	Arabidops
	19	1875 <td>54.0</td> <td>2256</td> <td>8</td> <td>AB054645</td> <td>Arabidops</td>	54.0	2256	8	AB054645	Arabidops
	20	1853 <td>53.3</td> <td>2640</td> <td>8</td> <td>AF355602</td> <td>Zea mays</td>	53.3	2640	8	AF355602	Zea mays
c	21	1847 <td>53.2</td> <td>144217</td> <td>8</td> <td>AC139748</td> <td>Medicago</td>	53.2	144217	8	AC139748	Medicago
	22	1844.5 <td>53.1</td> <td>2279</td> <td>6</td> <td>AR477456</td> <td>Sequence</td>	53.1	2279	6	AR477456	Sequence
	23	1841.5 <td>53.0</td> <td>2159</td> <td>8</td> <td>AB042322</td> <td>Arabidops</td>	53.0	2159	8	AB042322	Arabidops
	24	1836.5 <td>52.8</td> <td>2181</td> <td>8</td> <td>AK04831</td> <td>Oryza sat</td>	52.8	2181	8	AK04831	Oryza sat
	25	1833.5 <td>52.8</td> <td>2338</td> <td>8</td> <td>AK072809</td> <td>Oryza sat</td>	52.8	2338	8	AK072809	Oryza sat
	26	1830.5 <td>52.7</td> <td>2299</td> <td>8</td> <td>AJ704374</td> <td>Brassica</td>	52.7	2299	8	AJ704374	Brassica
	27	1817 <td>52.3</td> <td>1940</td> <td>6</td> <td>AX654768</td> <td>Sequence</td>	52.3	1940	6	AX654768	Sequence
	28	1811 <td>52.1</td> <td>2167</td> <td>8</td> <td>BA311388</td> <td>Brassica</td>	52.1	2167	8	BA311388	Brassica
	29	1809.5 <td>52.1</td> <td>2166</td> <td>8</td> <td>AJ633706</td> <td>Brassica</td>	52.1	2166	8	AJ633706	Brassica
	30	1807.5 <td>52.0</td> <td>1971</td> <td>6</td> <td>BD175021</td> <td>Gene enco</td>	52.0	1971	6	BD175021	Gene enco
	31	1807.5 <td>52.0</td> <td>1971</td> <td>8</td> <td>AB049624</td> <td>Arabidops</td>	52.0	1971	8	AB049624	Arabidops
	32	1807.5 <td>52.0</td> <td>2000</td> <td>6</td> <td>AX653563</td> <td>Sequence</td>	52.0	2000	6	AX653563	Sequence
	33	1799 <td>51.8</td> <td>2310</td> <td>8</td> <td>BT009249</td> <td>Triticum</td>	51.8	2310	8	BT009249	Triticum
	34	1799 <td>51.7</td> <td>2248</td> <td>8</td> <td>AR477465</td> <td>Sequence</td>	51.7	2248	8	AR477465	Sequence
	35	1796.5 <td>51.7</td> <td>2248</td> <td>8</td> <td>AF332959</td> <td>Lycopersi</td>	51.7	2248	8	AF332959	Lycopersi
	36	1793.5 <td>51.6</td> <td>2808</td> <td>8</td> <td>AK121195</td> <td>Oryza sat</td>	51.6	2808	8	AK121195	Oryza sat
	37	1791.5 <td>51.6</td> <td>2327</td> <td>8</td> <td>AF347613</td> <td>Lycopersi</td>	51.6	2327	8	AF347613	Lycopersi
	38	1789 <td>51.5</td> <td>2310</td> <td>8</td> <td>AB061739</td> <td>Arabidops</td>	51.5	2310	8	AB061739	Arabidops
	39	1777.5 <td>51.2</td> <td>1989</td> <td>8</td> <td>AX653919</td> <td>Sequence</td>	51.2	1989	8	AX653919	Sequence
	40	1777.5 <td>51.2</td> <td>2791</td> <td>8</td> <td>AK103007</td> <td>Oryza sat</td>	51.2	2791	8	AK103007	Oryza sat
	41	1773.5 <td>51.0</td> <td>1989</td> <td>8</td> <td>AF493792</td> <td>Oryza sat</td>	51.0	1989	8	AF493792	Oryza sat
	42	1771.5 <td>51.0</td> <td>2494</td> <td>8</td> <td>AK067270</td> <td>Oryza sat</td>	51.0	2494	8	AK067270	Oryza sat
	43	1770.5 <td>50.9</td> <td>2262</td> <td>8</td> <td>BA416460</td> <td>Brassica</td>	50.9	2262	8	BA416460	Brassica
	44	1767.5 <td>50.9</td> <td>2439</td> <td>8</td> <td>AF309643</td> <td>Solanum t</td>	50.9	2439	8	AF309643	Solanum t
	45	1760.5 <td>50.7</td> <td>2442</td> <td>8</td> <td>HVHVT1</td> <td>Horduem vul</td>	50.7	2442	8	HVHVT1	Horduem vul

ALIGNMENTS

RESULT 1
AR477464
LOCUS AR477464 2449 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 17 from patent US 6696292.
ACCESSION AR477464
VERSION AR477464.1 GI:47234922
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2449)
AUTHORS Allen,S.M., Falco,S.C. and Thorpe,C.J.
TITLE Genes encoding sulfate assimilation proteins
JOURNAL Patent: US 6696292-A 17 24-FEB-2004;
FEATURES Location/Qualifiers
source
1..2449
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 5.03e-268 Length: 2449
Score: 3475.00 Matches: 680
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-762-049-18 (1-680) x AR477464 (1-2449)	
QY	1 HisGluLeuAlaArgThrLeuSerTyrIleThrHisIleCysLeuLeuArgAsnThrIle 20
Db	2 CAGGAGTAGCTGCACATTAAGTTATATTAACATATTTGCTTGTAGAAATACTATT 61
QY	21 IleGluAspMetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArg 40
Db	62 ATTGAAGATATGGGAGGTAGATTATGAGTACCCTTTGGGCATGAACAACTTTTGAGA 121
QY	41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
Db	122 GTGCACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAGGTCTCTAAAGTAGTCTCTTG 181
QY	61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db	182 AAGGAGACTTCTTCCTGATGACCTTTGAGGAGTTCAAGAACAGCCAGCTTCCAG 241
QY	81 LysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThr 100
Db	242 AAGTTCAATGCTTGGCTTCAGTCTCTTCCCATTTTCGAATGGGCTCCCAAAATACACC 301
QY	101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
Db	302 TTTTCAGTCTCTGAAGCTGACCTCAGTGGCATCACCATCGCTAGCTTGGCCATTCTCT 361
QY	121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140
Db	362 CAGGGCATCAGTTATGCCAGCTCGCCACCTCCCTCCCAATCTTGAGACTATATTCGAGC 421
QY	141 PheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db	422 TTTATACCACTTGAATTTATCGATCATGGGTAGTCGAGGAGATTGGCAGTGGGACT 481
QY	161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
Db	482 GTGGCGGTGGATTCGCTTCTGATGGGTTCGATGTGAGTAATGCCGTGTATCCCAATGAA 541
QY	181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
Db	542 GACCCAAAGCTTTACCTCCACCTGGCTTTCACAGCTACATATTGCTGGTGTTCAG 601
QY	201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
Db	602 GCTGCTCTGGGTCTGTTAGTTGGGGTTCGATCGTGAATTTCTGTACATGCAACCAT 661
QY	221 IleGlyPheMetGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeu 240
Db	662 ATAGGGTTCTATGGGAGAGCAGCCGCTGGTGTCTGAGCAACTAAATCGATCTT 721
QY	241 GlyLeuGluHisPheThrHisGlyAlaAspIleSerValMetArgSerValPheThr 260
Db	722 GGCTTTGAGCATTTCCACCATGAGCTGATATCATATCAGTGATGGCTCTGTTTCACC 781
QY	261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280
Db	782 CAAACTCATGAGTGGAGGTGGGAAAGTGTCTGTAGGATGTCTTCATTTCTTCTCTC 841
QY	281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
Db	842 CTTAGCACAAAGATACCTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCG 901
QY	301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db	902 CCATTGACGTCCTTATATGGGAAGTCTCTTGGTTTATTTCACTCAGCCCGAGAGCAC 961
QY	321 GlyValGluValIleGlyLeuLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
Db	962 GGTGTTTGAAGTAGTAGAGAACTGAAGAAGGGTTTGAATCCACCATCACTCAAAATCTG 1021
QY	341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
Db	1022 GTATTTGTGTCGCTTACATGACTACAGCTGTCAAAACTGGCANTTGTCTTGGCATCAT 1081
QY	361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
Db	1082 TCACTTGGCGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTTACAATATT 1141
QY	381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
Db	1142 GATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACTGATAGTTGGTTCTTTCCACTCT 1201
QY	401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
Db	1202 TGTACTACCTCACAAAGAGACCATTTTCGGCTTCGGCTGTGAACCTATAACGCTGGATGCAAG 1261
QY	421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db	1262 ACAGCAGCTTCCAAACATATATATGTCACCTTGAGTAATGTTGACATTTGTTATTCCTGACA 1321
QY	441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db	1322 CCCTTGTTCATATCACATCCCTCGTGGTGTCTATCAGCTATTATCGTATCTGCAATGCTT 1381
QY	461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db	1382 GGACTCATAGATTATGAAGCAGCCATCCATCTATTATAAGTTGACAAATTTGACTTTGTG 1441
QY	481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db	1442 GTGTGATGAGTGATCATATTCGGCTGGTCTTTGGCAGCTTGGAAATTTGGCTTAGTCATA 1501
QY	501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
Db	1502 GCTATTGTAATATCTGTACTTCGGGTACTTCTATTATTGCAAGGCCAAGGACATTCGTT 1561
QY	521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db	1562 TTGGGCACATTTCCAAATTCGTGTATATACCGAAATGTTGAGCACTATCCAAATGCCAAA 1621
QY	541 HisValProGlyMetLeuLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db	1622 CATGTTCTCGGAATGCTAATTTCTAGAGATTGATGCCCAATTTACTTTGGCAATGCCAGC 1681
QY	561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluGluArgIleLysAlaThr 580
Db	1682 TATTTAAAGAGAAGGATCACAGGTGATTCATGAAGAAGAAAGAAATTAAGCTACA 1741
QY	581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db	1742 GGGGAGACTAGTTTGCAGTATGTTAATATGATGAGTGTCTGTGGAAACATTGATACA 1801
QY	601 SerGlyIleSerMetLeuGluValLysIleThrGluArgArgGluLeuGlnLeu 620
Db	1802 AGTGGAAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGCTACAGCTT 1861
QY	621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db	1862 GTTTTGGTCAATCTCTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAAAT 1921
QY	641 HisLeuGlyLysLysTrpIleTyrLeuThrValGluAlaValGlyAlaCysAsnPhe 660
Db	1922 CATTTAGGGAAGAAATGATCTATCTGACTGTTGAAGAGGCGGTTGGAGCATGCAACTTC 1981
QY	661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db	1982 AATCTACGTGCAAGCAAAACGAAACCCAAAGAAGATGAACAGAGGGTTGGAACTATGTG 2041
RESULT 2	
D89631	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
D89631	
Arabidopsis thaliana mRNA for sulfate transporter, complete cds.	
D89631	
Arabidopsis thaliana	
ATST1; sulfate transporter.	
Arabidopsis thaliana (thale cress)	
2045 bp mRNA linear PLN 14-FEB-2004	

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsiis.

REFERENCE

1 Yamaguchi, Y., Nakamura, T., Hatada, E., Koizumi, N. and Sano, H.
The electronic Plant Gene Register
Plant Physiol. 113 (4), 1463-1465 (1997)

MEDLINE 9

97267155@post.ac.jp: 145 477, 1405 1406 (1997)
2 (bases 1 to 2045)
Yamauchi, Y.
Direct Submission
Submitted (29-NOV-1996) Yube Yamauchi, Nara Institute of Science
and Technology; 9916-5 Takayama, Ikoma, Nara 630-01, Japan
(E-mail: yu-yama@bemailgate.aist-nara.ac.jp, Tel: 07437-2-5652,
Fax: 07437-2-5659)

FEATURES

```

Location/Qualifiers
1. 2045
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/tissue_type="above-ground parts"
/clone_lib="lambda ZAP II"
/ecot_vpe="Landsberg erecta"

```

gene

```
1. .2045
/aene="ATST1"
```

CD5

```

33. .2009
/genbank=1
/codon_start=1
/product="sulfate transporter"
/protein_id="BAA21657.1"
/db_xref="GI:2285885"
/translation="MCTEDYTFQQAEEELHRRHHTVEAPQEPFELKSLQYSVKETLFF
DPDFROFNKQASRKEVGLKYLPIFEPWAPNRLLFFKSDLIAGITIASLAIPQGIS
YAKLANLPILGLYSFVPLVAVIYAGSSRDLAVGTAVASILLTGAMLSKEVDAKDP
KLYLHLAFTATPAGVLEASLGIRLGFIVDFLSHATIVGPMGGAATVSLQQLGFI
KXKLTFTSDTIVLSVMKRSYTHQWVSGVLSGCCGFLFELLSTRYSIKKPKPFVWA
MAPLTSIVLGSLLIVYTHAEHQVQVIGDLKGNLPGUSDLLITFSYFMSTAVKTLGI
TGIIALAEAGVGRSFAMFNKYNLIDKEMFIAGMNNI VGSFTSCYLTGTPPPRSADV
NAGCKTAMSNIVMAIYATNLTPLFPHYTPVLVLSIIISAMLGILDIYQAAHLV
KYDFKDFLMSGAYGVGFSIGIYAVAIISARLLFVSPRTAVKGNIPNSMIY
RNTQYFPPSRTQPGILIEDAIPAYANASYLRIIRTWIDEEERVKSGESSLQYI
ILMDSAMVINDTSGISMMVEIKKVIDRRALKLNSLPKNGVVKLITRSKFIQDHLKSE
WMFITYVGEAVEACSYMIHTFKTEPASKRNPNNV"

```

ORIGIN

Alignment Scores:	1.4e-201	Length:	2045
Pred. No.:	2642.50	Matches:	501
Score:	89.16%	Conservative:	80
Percent Similarity:	76.02%	Mismatches:	75
Best Local Similarity:	76.04%	Indels:	3
Query Match:	8	Gaps:	3
DB:			

US-10-762-049-18 (1-680) x D89631 (1-2045)

24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43

22 ATCCCGCAGCCACCTTTCCTCAAGGAGCGGCAGGTTCACGCCAC

3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527

[illegible]

DD 73 ACGGTGGAGGCATCCAGATCCCTCAACCTTCTTGGAGTCACTTCAGTACCTCAGTGAAGGAA 15

Qy 63 ThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysphe 82

Db 153 ACTCTGTTCCAGACGACCTTTTAGACAAATTAGAACCAAAATGCATCAAGAAATTT 21

Qy 83 MetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPheGln 10

D_b 213 GTGTTAGGCGCTCAAAATACCTTCCTCCCGATTTCGAATGGGCACCAACGGTACAATCTCAAG 27

100

[illegible]

Db	1353	ATTGACTATCAAGCTGCCATCTCTCGAAGATTGACAAGTTCGACTTCCTCGTCTGC	1412
Qy	483	MetSerAlaTyrLeuGlyValValPheGlySerValGluLeuGlyLeuValIleAlaIle	502
Db	1413	ATGAGCGCCTACGTTGGGGTGGTATTTCGGCAGTGTAGAGATTGGACTCGTGGTACGGGTG	1472
Qy	503	ValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuGly	522
Db	1473	CGCATATCTATAGCGAGTGTGTTGCTGTTGTCGAGGCCAAACCTGCGGTGAAGGGA	1532
Qy	523	AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVal	542
Db	1533	ACATACCAACAGCATCTATAGGAACACTGAGCAGTACCCATCATCAAGAACCGTT	1592
Qy	543	ProGlyMetLeuLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu	562
Db	1593	CCTGGTATCTCATCTTGGAGATTGATGCTCCCATCTACTTTGCTAATGCCAGTTACTTG	1652
Qy	563	ArgGluArgIleThrArgTrrPleAspGluGluGluArgIleLysAlaThrGlyGlu	582
Db	1653	CGTGAAGAATCATAGGTGGATTGATGAAGAGAGAGAGAGTAAACAATCAGGAGAG	1712
Qy	583	ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly	602
Db	1713	AGCATGTTACAAATATATTACTCGATATATCTAGCTGTTGGTAAATATCGACACAGCGGT	1772
Qy	603	IleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLeu	622
Db	1773	ATTAGCATGATGCTGGAAATTAGAAAGTCATTGACAGGAGACGCTTAAGTTGGTATTG	1832
Qy	623	ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis	641
Db	1833	TCAATTCCTCAAAAGAGAGAGGTCGTGAAGAAATTAACCAAGATCCCAATTCATCGGTGATCAT	1892
Qy	642	LeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn	661
Db	1893	TGCGCAAGAGTGGATGTTCTTAAACGTAGGAGACGATGAGCGCTGTAGCTACATG	1952
Qy	662	LeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrrPheAsnVal	680
Db	1953	CTTCACACGTTTAAACCGAACCG---GCCTCCAAAACGAGCGCTTGGACACACGTA	2006
RESULT 3			
AY099873			
LOCUS	Arabidopsis thaliana	2266 bp	mRNA linear PLN 06-MAY-2002
DEFINITION	complete cds.		
ACCESSION	AY099873		
VERSION	AY099873.1	GI:20466813	
KEYWORDS	FLI_CDNA.		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana (thale cress)		
REFERENCE			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
	1 (bases 1 to 2266)		
	Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.		
TITLE	Submitted (24-APR-2002)		
JOURNAL	Direct Submission		
	Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA		
COMMENT	e-mail for correspondence: arab@sequence.stanford.edu		

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

1. .2266

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="RAF109-95-N16 (R25423)"

/ecotype="Columbia"

/note="This clone is in pBluescript"

1. .2266

/gene="At3g51895"

84. .2060

/gene="At3g51895"

/codon_start=1

/product="sulfate transporter ATST1"

/protein_id="AA020724.1"

/db_xref="GI:20466814"

/translation="MGTEBDYTPQGAELHRRHHTVEAPQPPFLKSLQYSVKETLFP
DDFFQFNQNASRKFVLGKFLPIFEWAPRYNLKPFKSDIAGITIASLAPQGIS
YAKLANLPILGLYSFVPLVYVLGSSRLDVAVTVALSLTGAMLSKEYDAEKDP
KLYLHLATPTATFAGVLEASLGIFRLGTFIVDFLSHATVFGMGAAATVVSLOQLKGI
GLKHTSDTIVSVMSRVSPOHEWRESVGLCGFLPFLSTRVFSIKKPFVWAA
MAPLTSVLGVYTHAHRHGVQVIGDLKGLNPLSGDLITSPYMSATAVKTGLI
TCIILALRGVAVGRSFAPKYNIDGNKEMIAFGMMNIVGTSFSCYLTGPPSRANV
YNAGCTAMSNIVMAIAVMFLTPILPIFYHPLVLSAIIASMLGLIDYQAALHLW
KYDFDFLVCMSAYVGVVGVSGVIGLWVAISIALLLFVSRPTKAVKGNTPNSMIY
RNTGQPSRTPVGLIILEDAPIYFANASYLRERIRIWIDEEERVKSGSSSQYI
ILDMSAVGNIDTSGISMMVEIKKVDRLAKLVLSNPKGEVVKLTRSKFIDGHLGKE
WMFLTVGEAVEACSVMYLTFTKTEPASKNEPNNV"

ORIGIN

Alignment Scores:

Pred. No.:	1,92e-201	Length:	2266
Score:	2641.50	Matches:	500
Percent Similarity:	88.16%	Conservative:	81
Best Local Similarity:	75.87%	Mismatches:	75
Query Match:	76.01%	Indels:	3
DB:	8	Gaps:	3

US-10-762-049-18 (1-680) x AY099873 (1-2266)

Qy	24	MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln	43
Db	84	ATGGGACGCGGAGGACTACACATTCCTCAAGGACGGAGGAGCTTCACGCGCCACACAC	143
Qy	44	---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu	62
Db	144	ACGGTGGAGGCTCCACACCTCAACCGTCTCTGAAGTCACTTCAGTACTCAGTGAAGNA	203
Qy	63	ThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe	82
Db	204	ACTCTGTTTCCAGACGACCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT	263
Qy	83	MetLeuGlyLeuGlnPhePhePheProIlePheGluTrrAlaProLysTyrThrPheGln	102
Db	264	GTGTTAGGCTCAAAATACTTCTCCCGATTTTCGAATGGGACCGACTACAAATCTCAAG	323
Qy	103	PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly	122
Db	324	TTCTTCAATCAGATCTCATCGCCGGAATCACCATCGTAGTCTGCCATCCCTCAGGCG	383
Qy	123	IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIle	142


```
Db 384 ATCAGTTACGCCAACCTTGCTAACTTGCCTCCCAATCTTGGCCCTTTATTCAGTTTGTGA 443
Qy 143 ProProLeuLeuIleValAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 444 CCGCCATTGGTATACCGCGGTCTAGGAGGTTCAAGGACTTAGCGGTGGAAACGGTTGCG 503
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAlaAspProAsnGluAlaAspPro 182
Db 504 GTTGCCTCTCTGTTCAGAGTGGCGATGCTGAGCAAGAGTGTGATGCTGAGAAAGATCCT 563
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 564 AGCTTTACCTTACCTTGTCTTTCACCGCCACATTTTTCGCGCGGTCTCGAAGCCTCT 623
Qy 203 LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleGly 222
Db 624 CTTGGAAATTTTCAGTTAGGTTTCATAGTCGACTTCTATCGCATGCAACGATAGTAGGA 683
Qy 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeu 242
Db 684 TTCATGGGAGGAGCAGCGCGTGGTGGAGTCTGCAACAGCTTAAGGGTATTTTCGGACTT 743
Qy 243 GluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThr 262
Db 744 AACATTTTCAGACTCTCCGATGTTATCTGTCAATGCTTCCGTTTCTCCCAACT 803
Qy 263 HisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeuLeuSer 282
Db 804 CACGAGTGGAGATGGGAAAGTGGGTTCTTGGATGGTCTTCTATCTTCTTCTCTCC 863
Qy 283 ThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeu 302
Db 864 ACCAGATATTTTCAGCATCAAGAACCAAAATCTTTTGGGTGGCGGATGGCTCTCTTG 923
Qy 303 ThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVal 322
Db 924 ACCTCAGTGATCTTGGAAAGTCTCTTGGTTTACTTCACTCACGCTGAGAGACATGGTGT 983
Qy 323 GluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPhe 342
Db 984 CAAGTGATAGGGACCTGAGAAAGGTTGAATCCACTCTCCGGTCTGATCTCATCTTT 1043
Qy 343 ValSerProTyrMetThrAlaValLysThrGlyIleValValGlyIleLeuSerLeu 362
Db 1044 ACTTCCCTTACATGTCCACAGCTGTCAAACTGGCCTCATCATCTGGCATCAITGGCTCTC 1103
Qy 363 AlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly 382
Db 1104 GCTGAAGGATAGCAGTGGGAGAGGTTTTCGATGTTCAAGAACTACAAACATAGACGGG 1163
Qy 383 AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402
Db 1164 AACAAAGATAGTAGGTTGGATGATGAAATCATCTGTTGGTTCCTTCATCTTGTATC 1223
Qy 403 LeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla 422
Db 1224 CTCACACCGGACCATTTTCAAGTTCGCGAGTGAATACAAACGCGGTTGCAAGACCGCA 1283
Qy 423 AlaSerAsnIleLeuMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeu 442
Db 1284 ATGTCCAAACATAGTAGTGGCGATTTGCGTATGTTTCACACTCCTCTTCCTCACACCGCTT 1343
Qy 443 PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu 462
Db 1344 TTTCACTACACACACTCTGCTGCTCTCTGCCATCATCATATCCGAAATGCTCGGACTC 1403
Qy 463 IleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCys 482
Db 1404 ATTGACTATCAAGCTGCCATCTCTCGAAAGTTCACAAGTTCGAACTTCTCTCGTCTGC 1463
Qy 483 MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle 502
```

```
Db 1464 ATCAGCGCCTACGTTGGGTCTGTTTCGCGAGTGTAGAGATTGGACTCGTCTAGCGGTG 1523
Qy 503 ValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGly 522
Db 1524 GCGATATCTATACGAGGTTGTTGCTGTTGTCGAGGCCAAAACTGCGGTGAAGGA 1583
Qy 523 AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaIleHisVal 542
Db 1584 AACATACCAAAACAGCATGATCTATAGAACACTGAGCAGTACCCTCATCAAGAACCGTT 1643
Qy 543 ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu 562
Db 1644 CTTGGTATTCTCATCTTGGAGATTGATCTCCCATCTACTTTGCTTAATGCCAGTACTTG 1703
Qy 563 ArgGluArgIleThrArgTrpIleAspGluGluGluGluArgIleLysAlaThrGlyGlu 582
Db 1704 CGTGAAGAAGTATCAAGTGGATTGATGAAGAGAGAGAGAGATTAAACAATCAGGAGAG 1763
Qy 583 ThrSerLeuGluTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly 602
Db 1764 AGCAGCTTACAATATATATCTGATATGTCAGCTGTTGGTAAATATCGACACAAACGGT 1823
Qy 603 IleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLeu 622
Db 1824 ATTAGCATGATGCTGAATTAAGAAGTCAATTCAGCAGGAGCGCTTAAGTTGGTATTG 1883
Qy 623 ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641
Db 1884 TCAATCCAAAAGAGAGAGGTCGTGAAGAAATTAACAGATCCAAATTCATCGGTGATCAT 1943
Qy 642 LeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661
Db 1944 TTGGGCAAGAGTGGATGTTCTTAACGCTAGGAGAACGATGGAGGCTTGTACTCATATG 2003
Qy 662 LeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 2004 CTTACACGTTTAAACCGAACCG---GCCTCCAAAACGAGCCTTGGAAACAACGTA 2057

RESULT 4
BNAS581745      2193 bp      mRNA      linear      PLN 02-SEP-2003
LOCUS           Brassica napus mRNA for sulphate transporter (bst3.1 gene).
DEFINITION      AJ581745
ACCESSION       AJ581745.1 GI:34481597
VERSION         bst3.1 gene; sulphate transporter.
KEYWORDS        Brassica napus (rape)
SOURCE          Brassica napus
ORGANISM        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE       1
AUTHORS         Buchner,P., Stuijver,E., Hawkesford,M.J. and de Kok,L.J.
TITLE           Analysis of the effect of sulphur nutrition and atmospheric H2S
                exposure on the expression of sulphate transporter genes in Curley
                kale
JOURNAL         Unpublished
REFERENCE       2 (bases 1 to 2193)
AUTHORS         Buchner,P.
TITLE           Direct Submission
JOURNAL         Submitted (01-SEP-2003) Buchner P., Agricultural and Environmental
                Division, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ,
                UNITED KINGDOM
FEATURES        Location/Qualifiers
                source             1..2193
                organism="Brassica napus"
                /mol_type="mRNA"
                /db_xref="taxon:3708"
                1..2193
                /gene="bst3.1"
                1..79
                /genes="bst3.1"
                80..2056
                CDS
```

/gene="bet3.1"		Db		647 GTGGATTTTCTATCGCATCGCAACGATAGTGGGATTCATGGAGGAGCAGCGCGTGGTG 706	
/evidence=experimental		Qy		232 CysLeuGlnGlnLeuLysSerIleleuGlyLeuGluHisPheThrHisGlyAlaAspIle 251	
/product="sulphate transporter"		Db		707 AGTCTGCAACAGCTTAAAGGTATTTTGGACATTAACATTTACAGAAACCCACTGACGTT 766	
/protein_id="CA546442.1"		Qy		252 IleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGluSerAlaVal 271	
/db_xref="GI:34481598"		Db		767 ATTCTGTGTCATCGGTTTCAGTCTTCTCAAACTCATCTGAGATGGAGATGGTGTT 826	
/translations="MGTEONTFPQGAEEPHRRHHAVEADPPQPFLLSKLSQYSVKETLFP DDPFPQKQNTSRQVVLGLKFLPILWEAPLVNFKLFXSDLIAGTITIASLAIPOGIS YAKLANLPFTLGLYSFVPLVAFVLGSSKDLAVGTAVGSLLTGAMLKEVDAEKDP KLVLHAFATFAGVLEASLGI FRLGFI VDFLSHATIVGPMGAAATVVSQQLKGF GLKFTETADVIVSMRSVSQTHQWRWESGLGCCFFLLSTRYSTRKKPFPWAA MAPLTSVLIGLIVYTHAERHGVQVIGDLKGLNPLSVDLVFTSPYMSALKTGLI TGIITLAEAGVRSFAMKFNKYNIDGNKEMIAFGMNI VGSFTSCVLTTPGFSRAVN FNAGCTAVSNI VMAIAVMTLLFHTPFPHYTPLVLSLI IMVAMGLDIYQAAIHLW KVDKPDFVCMASVAGVFGVSGEIGLVAVVISIARLLFLVSRPRTAVKGNIPNTMIY RNTDOYPSRI VPGLLILEIDAPIYANAGYLREIRTWVDEEDRIKASGNSLOVY ILDMASVGNIDTSGISMBEIRKIMDRREKLVIANPKGEVVKULTRSRFIGNLGSKE WMLPTVGEAVEACSFMLHTSKTEPASPKEEPPWNV"		Qy		272 LeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSerLysLysArgPro 291	
2057. .2180		Db		827 CTCGGCTGTTGTTCT 886	
/gene="bet3.1"		Qy		292 ArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleleuGlySerLeuLeu 311	
3' UTR		Db		887 AAAATTCTTTTGGGTAGCTGCGATGGCTCTTTGACCTCAGTCACTTCTTGGAAAGCCCTCTTG 946	
ORIGIN		Qy		312 ValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGly 331	
Alignment Scores:		Db		947 GGTACTTTCACTCACGCGGAGAGACATGGTGTTCAAGTGATAGGGGACCTGAGAAAGG 1006	
Pred. No.:	6.06e-198	Length:	2193	332 LeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaVal 351	
Score:	2597.50	Matches:	497	1007 TTGAACCCGCTCTCTGTGTCTGATCTTGTCTTCACTTGGCCCTACATGTCAACAGCTCTC 1066	
Percent Similarity:	86.27%	Conservative:	81	352 LysThrGlyValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSer 371	
Best Local Similarity:	74.18%	Mismatches:	87	1067 AAAACTGGCCCTCACTGGGATCATTACTCTTGCAGAAAGGATAGCTGTGGAAAGGAGC 1126	
Query Match:	74.75%	Indels:	5	372 PheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThr 391	
DB:	8	Gaps:	4	1127 TTTGCCATGTTCAAGAACTACAACTAGACGGTAACAAGAAATGATACGTTTGAAGTG 1186	
US-10-762-049-18 (1-680) x BNA581745 (1-2193)		Qy		392 MetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSer 411	
Qy		Db		1187 ATGAACATCGTTGGTTCTCTTCACTTGTACCTCAAACTGGACCGTTTTCAAGATCG 1246	
Qy		Qy		412 AlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAla 431	
Qy		Db		1247 GCCGTGAATCTCAATGCGGGTTGCAAGCTGCCGTGCCAATAGTAAATGCCCATGGC 1306	
Qy		Qy		432 ValMetLeuThrLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeu 451	
Qy		Db		1307 GTATGTTCCAGATTGCTCTTCCACAGCGCTTTTTCCTACACGCGCCCTCGTGCCTC 1366	
Qy		Qy		452 SerAlaIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeu 471	
Qy		Db		1367 TCCTCCATCATCGTGCATGCTCGACTCATCGACTATCAAGTCGCCATCCATCTT 1426	
Qy		Qy		472 PheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPhe 491	
Qy		Db		1427 TGGAAAGTTGACAAATTCGATTTCTCGTCTGCATGAGCGCCTACTTTGGGGCTGTGTTT 1486	
Qy		Qy		492 GlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeu 511	
Qy		Db		1487 GGTAGTGTGAGATCGGACTGTTGTCGAGTGGTAATATCGATAGCAAGGTGTGTGCTG 1546	
Qy		Qy		512 PheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArg 531	
Qy		Db		1547 TTCGTGTGCGGCCCGGAGCTGCGGTGAAAGGAAACATACCAACACCATGATCTATAGG 1606	
Qy		Qy		532 AsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAsp 551	
Qy		Db		1607 AACACTGCCAGATCTCTTACTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1666	
Qy		Qy		552 AlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAsp 571	
Qy		Db		1667 GCTCCCATCTACTTTGCCAACGCTGGTTACTTGGCGGAGAGNATCACAAAGGTGGTCGAT 1726	
Qy		Qy		572 GluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAsp 591	

```
Db 1727 GAAGAGGAGATAGGATCAAGAGCATCAGGAGGAATAGTTTACAATATGTTACTCGAT 1786
Qy 592 MetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluValLysLys 611
Db 1787 ATGCTGCTGTTGTAATATCGACACGAGTGGTATAGCATGATGAGGAAATTAAGAA 1846
Qy 612 IleThrGluArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLys 631
Db 1847 ATCATGGACAGAGAGAGTAAAGTTAGTATTCGCAATCCAAAGAGAGAGTGGTGAAG 1906
Qy 632 LysLeuAsnLysSerLysPhe---GlnAsnHisLeuGlyLysLysTrpIleThrLeuThr 650
Db 1907 AAATACACAGATCCAAATTCATCGGTGAAATTTGGGCAAGAGTGGATGTTCTTAACA 1966
Qy 651 ValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerLysThrAsnProLys 670
Db 1967 GTTGGAGAGCCGTGGAAGCTTGTAGTTTCATGCTTCACACGTCAAAGACCGAACCGCC 2026
Qy 671 LysAspGluThrGluGlyTrpAsnAsnVal 680
Db 2027 TCCAAAGAA---GAGCCTTGGAAACAGTT 2053

RESULT 5
AX506807
LOCUS AX506807 1977 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1502 from Patent W00216655.
ACCESSION AX506807
VERSION AX506807.1 GI:23388044
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Harper, J.F., Krops, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 1502 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source 1. .1977
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 7.67e-198 Length: 1977
Score: 2595.50 Matches: 498
Percent Similarity: 87.73% Conservative: 81
Best Local Similarity: 75.45% Mismatches: 77
Query Match: 74.69% Indels: 5
DB: 6 Gaps: 3
US-10-762-049-18 (1-680) x AX506807 (1-1977)
Qy 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
Db 1 ATGGGCAGGAGGACTACACATTCCTCTCAAGGAGCGGAGAGCTTCACCGCCGCCACCAC 60
Qy 44 ---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 61 ACGGTGGAGGCTCCACAACTCAACCGTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120
Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 121 ACTCTGTTCCAGACGACCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 180
Qy 83 MetLeuGlyLeuGlnPhePheProLeuPheGluTrpAlaProLysTyrThrPheGln 102
Db 181 GTGTTAGGCTCAAACTACTTCTCCCGATTTTCGAATGGGCACCAACGCTACAAATCTCAAG 240
```

```
Qy 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 241 TTCTTCAAATCAGATCTCATCGCGGAATCACCATCGTAGTCTCGCCATCCCTCAGGC 300
Qy 123 IleSerTyrAlaIleValLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIle 142
Db 301 ATCAGTTACGCCAAACTTGTAACTTGGCCCCCAATTTCTTGGCCTTTATTTCGAGTTTGA 360
Qy 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 361 CGCCCATTCGTATACCGGTGTAGGAGTTCAGAGGACTTAGCGGTGGGAACGGTTGCG 420
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
Db 421 GTTGCCTCTCTGTTCACAGTGCATGCTCAGCAAGAAAGTTGATCTCTCAGAAAGATCTCT 480
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 481 AAGCTTTTACCTTACCTTGTCTTTCACCGCCACTTTTTTCGCCGCGCTTCTCGAAGCCTCT 540
Qy 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGln 222
Db 541 CTGTGAAATTTTCAAGGTTAGGGTTTCATAGTGGACTTTCTATCGCATGCAACGATAGAG 600
Qy 222 YPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242
Db 601 ATTCATGGAGGAGCAGCGCGTGGTGTCTGCAACAGCTTAAGGGTATTTTCGGACT 660
Qy 242 uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnTh 262
Db 661 TAAACATTTTCAAGACTCTACCGATGTATCTCTGTATGTCGCTTCGTTTCTCCCAAC 720
Qy 262 rHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
Db 721 TCACGAGTGGAGATGGGAAGTGGCGTCTTGGATGGTCTTCTATTCTTCTCTCTC 780
Qy 282 rThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLe 302
Db 781 CACCAGATATTGAGCATCAAGAAACCAAAATCTTTTGGTGGCGCGCATGCTCTCTTT 840
Qy 302 uThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGlyLysHisGlyVa 322
Db 841 GACCTCAGTGTATCTTGGAAAGTCTCTTGGTTTACTTCTCACTCACTGAGAGACATGGTGT 900
Qy 322 lGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342
Db 901 TCAAGTATAGGGGACCTGAAGAAAGGTTGAATCCACTCTCCGGTCTCTGATCTCATCTT 960
Qy 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLe 362
Db 961 TACTTCCCTTACATGTCACAGCTGTCAAAACTGGCCTCATCACTGGCATCATCTCTCT 1020
Qy 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGln 382
Db 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTCAACATCAGACGG 1079
Qy 382 YAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy 402
Db 1080 GAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCCTTCACATCTTGTTA 1139
Qy 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422
Db 1140 CCTCAACACCGGACCATTTTCAAGTTCGCGAGTGAACACTCAACCGGGTTGCAAGACCG 1199
Qy 422 alaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheIleuThrProLe 442
Db 1200 AATGTCCAACTAGTAGCGGATTTGCGGTTATGTTTACACACTCTCTCTCACACCGCT 1259
Qy 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462
Db 1260 TTTTCTACTACACCACTCGTCTCTCTCTGTCATCATCATATCCGCANTGCTCGACT 1319
```

Qy	462	uileAspTyrGluAlaAlaIleHisLeuPheHisValAspLysPheAspPheValValCy	482
Db	1320	CAITGACTATCAAGCTGCCATCCATCTCTGGAAGTTGACAAAGTTCGACTTCTCGTCTG	1379
Qy	482	sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI	502
Db	1380	CATGAGCGCTACGTTGGGGTCTGATTTCGGCATGTAGAGATTGGACTCGTCTAGCGGT	1439
Qy	502	eValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuG	522
Db	1440	GGCGATATCTATAGCGAGGTGTGTCTGTTGTCTGAGGCCAAACATCCGGTGAAGGG	1499
Qy	522	yAenIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa	542
Db	1500	AAACATACCAACACAGCATGATCTATAGGAACACTGAGCAGTACCCCATCATCAAGAACCGT	1559
Qy	542	lProGlyMetLeuIleGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLe	562
Db	1560	TCTGTGTATCTCATCTTGAGATTGATGCTCCCATCTACTTTGCTTAATGCCAGTTACTT	1619
Qy	562	uArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyG	582
Db	1620	GCCTGAAGAAATCATAGGTGGATTGATGAGAGGAAGAGAGATTAAACAATCAGGAGA	1679
Qy	582	uThrSerLeuGlnTyrValIleIleAspMetSerAlaValAlGlyAsnIleAspThrSerG	602
Db	1680	GAGCAGCTTACAAAT	1739
Qy	602	yIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValle	622
Db	1740	TATTAGCATGATGGTGAATTAAGAAAGTCATTGACAGAGAGCGGTTAAAGTTGTATTT	1799
Qy	622	uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHi	641
Db	1800	GTCAATCCAAAGAGAGAGCTCTGAGAAATTAACAGATCCAAATTTATCGTGATCA	1859
Qy	641	sLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValAlGlyAlaCysAsnPheAs	661
Db	1860	TTTGGGCAAAAGAGTGGATGCTTAAACGGTAGGAGAAGCAGTGGAGGCTTGAGCTACAT	1919
Qy	661	nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal	680
Db	1920	GCTTCACAGTTTAAACCGAACCG---GCCTCCAAAACGAGCCTTGGAAACAACGTA	1974
RESULT	6		
AX651748			
LOCUS	AX651748	1977 bp	DNA linear PAT 22-MAR-2003
DEFINITION	Sequence 589 from Patent WO0300898.		
ACCESSION	AX651748		
VERSION	AX651748.1	GI:29154566	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
REFERENCE	1	Chang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M., Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.	
AUTHORS	Plant genes involved in defense against pathogens		
TITLE	Patent: WO 0300898-A 589 03-JAN-2003;		
JOURNAL	Syngenta Participations AG (CH)		
FEATURES	Location/Qualifiers		
source	1. .1977		
	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
ORIGIN			
Alignment Scores:			
Pred. No.:	7, 67e-198	Length:	1977
Score:	2595.50	Matches:	498
Percent Similarity:	87.73%	Conservative:	81

Best Local Similarity:	75.45%	Mismatches:	77
Query Match:	74.69%	Indels:	5
DB:	6	Gaps:	3
US-10-762-049-18 (1-680) x AX651748 (1-1977)			
Qy	24	MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln	43
Db	1	ATGGGCACGGAGGACTACACATTCCTCAAGGAGCGGAGGAGCTTCACCGCGCCACCAC	60
Qy	44	---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu	62
Db	61	ACGGTGGAGGCTCCCAACCTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA	120
Qy	63	ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe	82
Db	121	ACTCTGTTTCCAGACGACCCCTTTAGACAAATTTAGAACCAAAATGCATCAGAAAAATTT	180
Qy	83	MetLeuGlyLeuGlnPhePhePheProIlePheGluTyrAlaProLysTyrThrPheGln	102
Db	181	GTGTTAGGCTCAAATACTTCTCCGGAATTTTGAATGGGCACCGCTACAATCTCAAG	240
Qy	103	PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly	122
Db	241	TTCTTCAAATCAGATCTCATCCCGGAATFCACCATCGCTAGTCTCGCCATCCCTCAGGCG	300
Qy	123	IleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerPheIle	142
Db	301	ATCAGTTACGCCAAACTTGTCTAACTTGGCCCCCAATCTTGGCTTTATTCGAGTTTGT	360
Qy	143	ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla	162
Db	361	CGCCATTGGTATACGGGTGTAGGGAGTTCAAGGAGCTTAGCGGTGGGACGGTTGCG	420
Qy	163	ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro	182
Db	421	GTTCGCTCTCTTTCACAGGTCGATGCTGACCAAGAGTTGATGCTCAGAAAGATCT	480
Qy	183	LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla	202
Db	481	AGCTTTACCTTCACCTTGTCTTCCCGGCATTTTTTCGCCGGGCTTCCGAGGCTCT	540
Qy	203	LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleG	222
Db	541	CTTGAATTTTCAGGTTAGGTTTCATAGTGACTTCTATCGCATGCAACGATAGTAGG	600
Qy	222	yPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeuGlyLe	242
Db	601	ATTTCATGGGAGGAGCAGCGCGTGTGAGTCTGCAACAGCTTAAGGGTATTTTCGGACT	660
Qy	242	uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnTh	262
Db	661	TAAACATTTACAGACTCTACCGATGTTATCTCTGTCACTCGCTTCGTTTTCTCCAAAC	720
Qy	262	rHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe	282
Db	721	TCACGAGTGGAGATGGGAAAGTGGGTTCTTGATGTGGTTTTCTATTCTTCTCTCTC	780
Qy	282	rThrArgTyrPheSerLysLysArgProArgPheThrValSerAlaMetAlaProle	302
Db	781	CACCAAGATTTTGGAGCATCAAGAAACCAAAATCTTTTGGGTGGCGGCGATGGCTCTT	840
Qy	302	uThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVa	322
Db	841	GACCTCAGTATCTTGGAGTCTCTTGGTTACTTCTACTCAGCTGAGAGACATGGTCT	900
Qy	322	lGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh	342
Db	901	TCAGTGTAGGGGACCTGAGAAAGGGTTGAATCCACTCTCCGGTTCTGTATCTCATCTT	960
Qy	342	eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLe	362
Db	961	TACTTCCCCCTTACATGTCCACAGCTGTCAAAACTGGCCTCATCATCTGGCATCATGTCT	1020

Qy 362 uAlaGluGlylleAlaValAlglyArgSerPheAlaMetTyrIysAsnTyrAsnIleAspG1 382
Db 1021 CGC-GAAGAGTATACAGTGGGAGGAGTTTGGCATGTTCAAGAACTACACATAGACGG 1079
Qy 382 YAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy 402
Db 1080 GAACAAAGATGATAGCGTTTGGATGATGAACATCGTTGGTCTCTTCAACATCTTGTTA 1139
Qy 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysIysThrAl 422
Db 1140 CCTCACACCGGACCATTTTCAAGTCGCGAGTGAACACTACACCGCGGTTTGAAGACCGC 1199
Qy 422 aAlaSerAsnIleMetSerLeuAlaValMetLeuThrIleuLeuPheLeuThrProLe 442
Db 1200 AATGTCACACATAGTATGCGGATGCGGTATGTTTCACACTCCTCTCTCCACCGCT 1259
Qy 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462
Db 1260 TTTTCACTACACACCATCGTCCTCTCTGCCATCATCATATCCGCATGTCGGACT 1319
Qy 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482
Db 1320 CATTGACTATCAAGTCGCCATCATCTCTCGAAAGTTGACAAAGTTCGACTTCTCGTCTG 1379
Qy 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI 502
Db 1380 CATGAGCGCTACGTTGGGGTCGTATTCGGCAGTGTAGATATGGACTCGTCTAGCGGT 1439
Qy 502 eValIleSerValLeuArgValLeuLeuPheIleAlaAatgProArgThrPheValLeuG1 522
Db 1440 GGGATATCTATACGAGGTGTGCTGTTTGTGCGAGCCAAATCTCGGTGAAGGG 1499
Qy 522 YAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa 542
Db 1500 AAACATACCAACAGCATGATCTATAGAACACTGAGCAGTACCATCATCAAGAACCGT 1559
Qy 542 lProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLe 562
Db 1560 TCCTGGTATTCTCATCTTGAGATTCATGCTCCCATCTACTTGTGTAATGCCAGTTACTT 1619
Qy 562 uArgGluAatgIleThrArgTrpIleAspGluGluGluGluArgIleLysAlaThrGlyG1 582
Db 1620 GCGTGAAGAAATCATAGGTGGATGTATGAAGAGGAGAGAGAGTTAAACATCAGGAGA 1679
Qy 582 uThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerG1 602
Db 1680 GAGCAGCTTACATATATTACTCGATATGTCAGCTGTTGGTAAATATCGACACAGCGG 1739
Qy 602 YIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLe 622
Db 1740 TATTAGCATGATGTTGGAATTAAGAAAGTCATTTCAGCAGAGAGCGTTAAAGTTGGTATT 1799
Qy 622 uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHi 641
Db 1800 GTCAATATCCAAAGAGAGAGTGTGAAGAAATTAACCGATGCCAATTCATCGGTGATCA 1859
Qy 641 sLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValAlaGlyAlaCysAsnPheAs 661
Db 1860 TTTGGCAAGAGTGGATGTTCTTAACGGTAGGAGAGCAGTGGAGGCTTGTAGCTACAT 1919
Qy 661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrTrpAsnVal 680
Db 1920 GCTTTCACACGCTTTAAACACCGCG---GCCCTCAAAAACGAGCGCTTGAAGAACGTA 1974

RESULT 7
AX653455
LOCUS AX653455 1974 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 3325 from Patent WO03000898.
ACCESSION AX653455
VERSION AX653455.1 GI:29156269
KEYWORDS Oryza sativa

ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 3325 03-JAN-2003;
SYNGENTA PARTICIPATIONS AG (CH)
FEATURES
Location/Qualifiers
source 1..1974
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Alignment Scores: 1.17e-176 Length: 1974
Pred. No.: 2330.50 Matches: 441
Score: 83.84% Conservative: 83
Percent Similarity: 70.56% Mismatches: 100
Best Local Similarity: 67.06% Indels: 1
Query Match: 6 Gaps: 1
DB: 6
US-10-762-049-18 (1-680) x AX653455 (1-1974)
Qy 43 GlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 52 AGGTCGCCGATGCGCGCGGAGCGTTCCTGGAGACGCTGGGGGGAACATGAAGGAG 111
Qy 63 ThrPhePheProAspAspProLeuArgGlnPhe---LysAsnLysProAlaSerLysLys 81
Db 112 ACATTCCTCGCGACACCGTTCAGGTCGTGTCGCGGAGCGCGGTGTCGCGGCGCGC 171
Qy 82 PheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThrPhe 101
Db 172 GCGCGCGCGCGCTCCGGTACGTGTTCCCGTTCATGGAGTGGCGCGCTGTACACCTC 231
Qy 102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db 232 GGCACCTCAAGTCCGACCTCATCGCGCGCATCACCATTGCCAGCTCGCCATCCCCAG 291
Qy 122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPhe 141
Db 292 GGCATCAGCTAGCCAAAGCTCGCAACCTCCCTCCGCTCTCGGCTCTATTCCAGCTTC 351
Qy 142 IleProProIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db 352 GTGCCCGCGCTGTGTACGCGATGATGGGAGCTCGAGGACCTGGCGGTGGGAGCGGTG 411
Qy 162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db 412 GCGGTGCGCTGCTGCTGATCGGTGATGCTGAGCGAGGAGTGTCCGCGCGGAGGAC 471
Qy 182 ProLysLeuTyrIleHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db 472 CCGGCGCTGTACGTGACGCTCGCGCTCACCGCCACCTCTTCTCCGCGCGGTTCAGGCG 531
Qy 202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db 532 CTGCTGGGGGTCTTGGAGCTGGGGTTCATCGTGGACTTCTGTCCGACCGCCATCGTC 591
Qy 222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db 592 GGGTTCATGGCGCGCGCCCGCCCGTGGTGTGCTGTCAGCAGCTCAAGGGCATGTTCCGC 651
Qy 242 LeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln 261
Db 652 CTCGACCACTTCACACCGCCACCGACCTCGTCTCCGTTCATGAGCTCCGTCTTCTCCCA 711
Qy 262 ThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeuLeu 281

```
Db 712 ACCACCTCTGGCGATGGGAGACGCTGTCATANGGCGTGGCGCTTCCTCTTCCTCCTC 771
Qy 282 SerThrArgTyrPheSerLysLysArgProArgPhePheTyrPheValSerAlaMetAlaPro 301
Db 772 ATCACCGCTTCTTACGACAGAGGACCAAGTTCTTCTGGGTATCTGCAGCTGGCCCA 831
Qy 302 LeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGly 321
Db 832 TTGGCGTCTGTATCATCGGAGGACCTCTGTTGTACTCATCATCTGAAACCATGCG 891
Qy 322 ValGluValIleGlyGluLeuLysLysGlyLeuAsnProSerLeuThrAsnLeuVal 341
Db 892 ATTCAGTGTATGTTACCTGGAAGAGGCGCTGAACCCACCGTCTCCCAACAGCTGAAC 951
Qy 342 PheValSerProTyrMetThrAlaValLysThrGlyLeuValValGlyLysIleSer 361
Db 952 TTCTCTGCGCGTACATGATGCTGGCCCTGGAACCGGATCATCTACTGCGTCAATTGCC 1011
Qy 362 LeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAsp 381
Db 1012 CTGCTGAGGGGATTCGCGTAGGAAGGAGCTTTGCAATGTTCAAGAAATTACCACATCGAC 1071
Qy 382 GlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCys 401
Db 1072 GGAACACAGGAGATGATCGCGTTTGAACGATGAACATTTGGGGATCACTACCTCTTGC 1131
Qy 402 TyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThr 421
Db 1132 TACCTCACACCGCTCGTTCGAGTTCGCGCTCAATCATCAACCCCGCTGCAGACG 1191
Qy 422 AlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrPro 441
Db 1192 GCGATGTCAACGATGATCGTGGTGGCGGTGATGATCAGCTGCTGTTCTTGACGCG 1251
Qy 442 LeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGly 461
Db 1252 CTGTTCCACTACACGCGCTGCTGGTGGTGTCTGGCGATCATCATGTCGCGCATGCTGGG 1311
Qy 462 LeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValVal 481
Db 1312 CTGATCGACTACCCGCGCGCTCCACTGTGGCAGGTGCACNAGGTGCACCTTCGCTC 1371
Qy 482 CysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla 501
Db 1372 TGCGTCGGCGCTTACCTCGCGCTGCTCTTGGCAGCGTGCAGATCGCTCGTGGCGC 1431
Qy 502 IleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeu 521
Db 1432 GTCGGATCTCCATCTCCGCGTGTCTGTTGTGGCGCGCGGAGGACGAGTGTCTC 1491
Qy 522 GlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHis 541
Db 1492 GGGAACTATCCCNACAGATGATCTACCGCGGATGACAGTACACCCCGCGCAGAGG 1551
Qy 542 ValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyr 561
Db 1552 GTCCCCGCGTGTCTGTCTCGCGTGCAGCTCGCCCATCTACTTCCACCAACGCCAGCTAC 1611
Qy 562 LeuArgGluArgIleThrArgThrIleAspGluGluGluArgIleLysAlaThrGly 581
Db 1612 CTGCGTGAAGGATTCGCGCGGTGGATTCGACGACGAGGAGGAGGATGTCGAAGGAGG 1671
Qy 582 GluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSer 601
Db 1672 GAGATGGGCATTCAGTACGTCTCGATGGGTGGGTGGAGTCCGAGCATCCACACCAGC 1731
Qy 602 GlyIleSerMetLeuGluGluLysLysIleThrGluArgArgGluLeuGlnLeuVal 621
Db 1732 GGGACGAGCATGTGATGAATCACTCAGGAAGACCCCTGGACAGGAGGCGCTTCAGATTGTG 1791
Qy 622 LeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHis 641
Db 1792 TTGGCGAACCCCGGGGAGGAGATCATGAAGAGTTGGACAGCTCCAGCTTCCAGGTGTTGAGGCG 1851
```

```
Qy 642 LeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661
Db 1852 ATCGGCCCATGATGATCTTCCCGACGGTGGCGAGCGGTGGCGAGTGCACACTTCG 1911
Qy 662 LeuArgAlaSerLys 666
Db 1912 ATGCACCTCGCAGAAG 1926
RESULT 8
AK066932
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J013093F02, full
insert sequence.
ACCESSION
AK066932
VERSION
AK066932.1 GI:32976950
FEATURES
FLJ_CDNA; CAP trapper.
SOURCE
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team:,
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shisniki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group:, Ootomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kohmaru,C., Kurotsaki,T., Kusumegi,T., Li,C., Lu,M.,
Saio,K., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
JOURNAL
Science 301 (5631), 376-379 (2003)
MEDLINE
22752273
PUBMED
12869764
REFERENCE
2 (bases 1 to 2327)
AUTHORS
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,T., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurotsaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Ootomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., and
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kanondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchienias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
```

COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shieshiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawanata, M., Kobayashi, M., Kodama, T., Kurosaki, R., Nukune, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nishikawa, T., Kikuma, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ota, Y., Saitho, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasakura, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

```
FEATURES
source
Location/Qualifiers
1. .2327
organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J01309P02"
```

ORIGIN

Alignment Scores:

Argument Scores:		
Pred. No.:	1.44e-176	2327
Score:	2330.50	Matches: 441
Percent Similarity:	83.84%	Conservative: 83
Best Local Similarity:	70.56%	Mismatches: 100
Query Match:	67.06%	Indels: 1
DB:	8	Gaps: 1

US-10-762-049-18 (1-680) X AK066932 (1-2327)

Qy	43	GlnValGluValIprProProGlnProPhePheIysSerLeuLysTyrSerLeuLysGlu	62
Db	163	AGGGTSCCGATGCGGCGCGAAGCCGTTCTGGAGACGCTGGGGGGGAACATGAAGGAG	222
Qy	63	ThrPhePheProAspAspProLeuArgGlnPhe---LysAsnLysProAlaSerLysLys	81
Db	223	ACATTCTCGCGACACCGCTTCAGGGTGGTGGCGCGGAGCGCGGTGGCGCGCGC	282
Qy	82	PheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThrPhe	101
Db	283	CGCGCGCGCGCGCTCCGTCACGTGTTCGGCTTCATGGAGTGGCGCGCGTCGACACCC	342
Qy	102	GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln	121
Db	343	GGCACCCCTCAAGTCGACCTCATCGCGGCATCACATTGCAGCGCTGCCCATCCCCAG	402
Qy	122	GlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuLeuGlyLeuTyrSerSerPhe	141
Db	403	GGCATCAGCTACGCCAAGCTCCCAACCTCCCTCCCGTCCTCGGCGCTCATTCGAGATT	462
Qy	142	IleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal	161
Db	463	GTCCCGCCGCTGGTGTTACCGATGATGGGAGCTCGAGGACCTGGCGGTGGGACCGTG	522
Qy	162	AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp	181

Db	523	CGGCTGGCGTCGTGCTGATCGGGTCATGCTGAGCGAGGAGGTGTTCGGCGCGGAGGAC	582
Qy	182	ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla	201
Db	583	CGGCGCGTGTATCTGCACGTCGCGCTCACCGCCACCTTCTTCGCGCGGTGTTCACGGCG	642
Qy	202	AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle	221
Db	643	CTGCTGGGGGTCTTGAGGCTGGGGTTTCATCGTGGACTTCTGTGCGACGCCACCATGCTC	702
Qy	222	GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly	241
Db	703	GGGTTCATGGCGCGCGCCGACCGTGGTGCTGTCGACGAGCTCAAGGCGATGTCGGC	762
Qy	242	LeuGlnHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln	261
Db	763	CTCGACCACTTACCACCGCCACCGACCTCGTCTCGTCAATGAGCTCCGTCTTCTCCCAA	822
Qy	262	ThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeu	281
Db	823	ACCCACCTCTGGGGATGGGAGCGCTCATGGGCTGGGCTTCTCTTCTTCTCTCTCTC	882
Qy	282	SerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaPro	301
Db	883	ATCACCCGCTTCTTCAGCAAGAGAGCGCAAGGTTCTTCTGGGTATCTGCAGCTGGCCCA	942
Qy	302	LeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGly	321
Db	943	TTGGCGTCTGCATCATCGGAGGCTCTCTGGTGTACTCACACATGCTGAAACCATGGC	1002
Qy	322	ValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuVal	341
Db	1003	ATTCAAGTGATGGTTACCTGAAGAAGGGGCTTGAACCCACCGCTGCGCACCAAGCCTGAAC	1062
Qy	342	PheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSer	361
Db	1063	TTCTCGTCGGCGTACATGATCTGGCCCTGGAACCGGATCATCACTGGCGTCAATGGC	1122
Qy	362	LeuAlaGluGlyIleAlaValAlGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAsp	381
Db	1123	CTCGCTGAGGGGATTCGCTAGGAGGAGCTTTTGCAATGTTCAAGAATTACACACATCGAC	1182
Qy	382	GlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCys	401
Db	1183	GGAAACAAGAGATGATCGCGTTTGGAAACCATGAACATTTGTGGGATCACTCACCTCTTGC	1242
Qy	402	TyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThr	421
Db	1243	TACCTCACCAACCGGTTCGTCTCGAGGTGGCGGTCAACTTACAAACCGCGCTGCAAGACG	1302
Qy	422	AlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrPro	441
Db	1303	CGGATGTCGAACGTGATCATGTCGGTGGCGGTGATGATCAACGCTGTGTCTTGTGACGCCG	1362
Qy	442	LeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGly	461
Db	1363	CTGTTCCACTACACGCGCTGGTGGTCTCTCGCGCATCATCATGTCTGGCGATGCTGGGG	1422
Qy	462	LeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValVal	481
Db	1423	CTGATCACTACCCGGCGCGCGTCCACCTGTGGCAGGTGGACAAGGTGGACTTCTCGCTC	1482
Qy	482	CysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla	501
Db	1483	TGCCTCGGCGCCCTACCTTCGGCGTCTCTTCGGCAGAGGTTCAGATTCGGCTCTGTCGTGCC	1542
Qy	502	IleValIleSerValLeuArgValLeuLeuPheIleAlaAlaArgProArgThrPheValLeu	521
Db	1543	GTCGGGATCTCCATCTCTCGCGTCTGCTCTTCTGTGGCGCGCGGAGGACGCGGTGCTC	1602
Qy	522	GlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHis	541
Db	1603	GGGAACATCCCCAACGATGATCTACCGGCGGATGGACCAAGTACACCGCGCGGAGG	1662

Db	907	GGTATCCAAATTATTCGGAGAACTAAAGAAAGGATAAATCCACCGTCAATCACACATTTG	966
Qy	341	ValPheValSerProTyrMetThrThraAlaValLysThrGlyLeValValGlyLeIle	360
Db	967	GTITTTACGCCACCTTATGATGTTGGCTCTCAAGTTGGCATTTCACTGGAGTCATA	1026
Qy	361	SerLeuAlaGluGlyLeAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle	380
Db	1027	GCTCTGCTGAGATATAGCAGTGGGGAGAGCTTTGCCAATGTACAGAAATTAACAATATA	1086
Qy	381	AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer	400
Db	1087	GATGGGAACAAGAGATAGATGCTTTGGGATGATGAACATTTCTGGTCTCTTCTCTCT	1146
Qy	401	CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys	420
Db	1147	TGCTATCTCACTACCGGCCATTTTCACGTTCCGGCGTGAACATACACGCGGGTTGCAA	1206
Qy	421	ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr	440
Db	1207	ACGGCGTTATCAACGTTGGTGATGGCGTTCGAGTGGCGGTGACGTTCGTCTTTAACG	1266
Qy	441	ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu	460
Db	1267	CCGTTATTTCTTACACGCCATTAGTGTCTGTCGTCGATCATATCGTCGCATGCTA	1326
Qy	461	GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal	480
Db	1327	GGCCTAGTCAGTACGACGAGCAGCATTCATCTTTGGAAATCGACAAAGTTCCGATTTCTTC	1386
Qy	481	ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle	500
Db	1387	GTITGTCCTCTCTACTTGGGGCTGTGTTTGGTACCATCGAAATTTGGTCTCATCCCTC	1446
Qy	501	AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal	520
Db	1447	TCGCTGGGAATACAGTAGAGGTAGTGTCTGTCGGAGACCAAGATTATGTA	1506
Qy	521	LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys	540
Db	1507	ATGGGAACACATTCAAAACCTCAGAAATATATAGGAACATTCAGCATTTACCCCTCAAGCCATC	1566
Qy	541	HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer	560
Db	1567	ACTCGCTCTTCTCTCTCATCTCCACATCGATGGTCCCTATTACTTCGCGCAACTTACC	1626
Qy	561	TyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThr	580
Db	1627	TACTTCGGAGACAGATTGGAAGTGGATCGATGAAGAGAGATAAATTGAGAACTAGC	1686
Qy	581	GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr	600
Db	1687	GGAGATATCAGTCTACAATATATTGACTCGACATGAGTCTGTGGGAAACATCGACACG	1746
Qy	601	SerGlyLeSerMetLeuGluGluValLysLysIleThrLeuArgArgGluLeuGlnLeu	620
Db	1747	AGCGGTATAGCATGCTTGAGGAACCTCAACAAATCTCTGGGACGTAGAGAGCTTAAGCTA	1806
Qy	621	ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn	640
Db	1807	GTGATAGCGNACCGGAGCAGAGGTGATGAAGAGCTAAGCAAGTCCACCTTCATTGAG	1866
Qy	641	HisLeuGlyLysLysTrpIleTyrLeuThrValGluAlaValGlyAlaCysAsnPhe	660
Db	1867	AGCATTTGGTAAAGACGATTTATCTCACGGTGGCAGAGCCGTCGACGTTTCGATTTTC	1926
Qy	661	AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnVal	680
Db	1927	ATGCTCCACACGGCTAAACCCGATCTCGCCGGTACCGGAA-----TTTAAACAGTT	1977
RESULT 10			
AJ601439			
LOCUS			

DEFINITION	Brassica oleracea mRNA for sulfate transporter (LST3.2 gene).
ACCESSION	AJ601439
VERSION	AJ601439.1 GI:37998857
KEYWORDS	LST3.2 gene; sulfate transporter.
SOURCE	Brassica oleracea var. acephala (kale)
ORGANISM	Brassica oleracea var. acephala Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE	1
AUTHORS	Buchner,P., Stuijver,E.E., Hawkesford,M.J. and de Kok,L.J.
TITLE	Analysis of the effect of sulphur nutrition and atmospheric H2S exposure on the expression of sulphate transporter genes in Curley kale
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2219)
AUTHORS	Buchner,P.
TITLE	Direct Submission
JOURNAL	Submitted (24-OCT-2003) Buchner P., Agricultural and Environmental Division, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UNITED KINGDOM
FEATURES	Location/Qualifiers
source	1..2219 /organism="Brassica oleracea var. acephala" /mol_type="mRNA" /variety="acephala" /db_xref="taxon:3713"
gene	1..2219 /gene="LST3.2"
5' UTR	1..71 /gene="LST3.2" /evidence=experimental
CDS	72..2024 /gene="LST3.2" /function="involved in plasma membrane sulfate transport" /product="sulfate transporter" /protein_id="CAE53112.1" /db_xref="GI:37998858"
5' UTR	2025..2219 /gene="LST3.2" /evidence=experimental
ORIGIN	
Alignment Scores:	
Pred. No.:	4,06e-172 Length: 2219
Score:	2274.50 Matches: 428
Percent Similarity:	81.50% Conservative: 105
Best Local Similarity:	65.44% Mismatches: 108
Query Match:	65.45% Indels: 13
DB:	8 Gaps: 4
US-10-762-049-18 (1-680) x AJ601439 (1-2219)	
Qy	30 GluTyrProLeuGlyMetAsnAsnPheGluArgValHisGlnValGluValProProPro 49
Db	90 CAGTAGACCC-----CAAGTAGAAACCCCTCCACCA 119
Qy	50 GlnProPhePheLysSerLeuLysTyrSerLeuLysGluThrPheProAspAspPro 69
Db	120 CAACCTTTCTTGAATCACTAAAGAACACTCTCAATGAGATCCTTTTCCCGACGATCG 179

```
QY 70 LeuArgGlnPheLysAsnLys----ProAlaSerLysLysPheMetLeuGlyLeuGlnPhe 88
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
180 TTCGAAATAATCAGAACGAATCGAAAAACGTCGAAAAAGATCGAGTTAGGTTGAGACAC 239

QY 89 PhePheProIlePheGluTrpAlaProLysTyrThrPheGlnPheLysAlaAspLeu 108
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
240 GTGTTCCCGATCTTGGAAATGGGTCGTGGTTTACAATTTGAGACTACTTGAATCAGACGTT 299

QY 109 IleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeu 128
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
300 ATCTCAGGATCAGATAGCAAGTCTTGCTATTCTCAAGGATAGTAGTATGCTCAACTC 359

QY 129 AlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIleProProLeuIleTyrAla 148
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
360 GCTAAATCTCCCTCAATCTTGTGCTTTACTCGAGCTTCGTGCGCGGATGGTATACGCG 419

QY 149 MetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuMet 168
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
420 ATAAATGGGAAGTTCAAAAGATTTAGCGGTGGCGACGCTAGCGGTGCGTCTGTTAAACG 479

QY 169 GlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeu 188
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
480 GCGCAATGCTTGGAAAGAGGTAAGCGCGGTGGAGAATCCAAAACCTTTACCTTCACTTA 539

QY 189 AlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeu 208
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
540 GCTTTTCACTCCACTTCTTCGCGGACTCATGCAAACTGCTTGGCTTTCACGGTTA 599

QY 209 GlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAla 228
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
600 GGATTCGTGGTGAGATATTGTCCGACGCGCGCATCATAGGGTTTCATGCTGGAGCAGCG 659

QY 229 ThrValValCysLeuGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGly 248
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
660 ACGTGCTTGTCTACAAACACTGAAGGATTACTTGGTCTCAGTCACCTTCACTCCCAATCC 719

QY 249 AlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlnTrpArgTrpGlu 268
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
720 ACTGATGTCGTCTCCGCTCTCTCGTTCATCTTTCTCAGTCTCCTGTGGAGATGGAG 779

QY 269 SerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSerLys 288
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
780 AGTGGGTACTAGGGTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 839

QY 289 LysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGly 308
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
840 AAGAGACCAAGCTATTCTGGATATCTGCAATGGCTCCACTTGTCTGTCAATTTTGGGA 899

QY 309 SerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeu 328
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
900 AGCTCTTCTGTTACTTCTCTCCATGCCAGTTCATGCCATCCAAATTAATGGAGATTA 959

QY 329 LysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThr 348
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
960 GAGAAAGGATTAATCCACGTCATCACACACTCTGTTTTCACGTCACCTTATGTCATG 1019

QY 349 ThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGlyIleAlaVal 368
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1020 CTGGCTCTTAAACCCGGCATTAATCACTGGAGTCTAGCTTGTGCTGAAGGAATTCAGTG 1079

QY 369 GlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAla 388
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1080 GGGAGAAGCTTTCATATGATAAGATTACACATAGATGGTAACAAAGATGATTGCA 1139

QY 389 IleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPhe 408
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1140 TTTGGGATGATGAACATTTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1199

QY 409 SerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleMet 428
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1200 TCACGTTCGCGGTGAACCTACACCGCCAGTTGCAAAACCGCGGTATCAAAACGTGGTGATG 1259
```

```
QY 429 SerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeu 448
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1260 CGGTTTCAGTAGCAGTAACGTTACTGTCTTAAACGCCATTATCTTTTATACGCCGTTA 1319

QY 449 ValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAla 468
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1320 GTAGTTCTGTCTCGATCATCATCGCAGCCATGCTTGGACTCGTCACTATGAAGCAGCC 1379

QY 469 IleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGly 488
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1380 ATGAATCTTTGAGACTCGACAAGTTCGATTCTTCGTATGTCCTCTCTCTCTCTCTCTCG 1439

QY 489 ValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArg 508
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1440 GTCGTTTTTGGTACCATTCGAGATTGGTCTCTCATCTCTCTGTGGGAATATCAGTCTCT 1499

QY 509 ValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerVal 528
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1500 TTACTGTGTTCTCGTGGGAAGACCAAAGATTTATGTAATGGGAACATTCAGAACACAGAG 1559

QY 529 IleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeu 548
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1560 CCTATAGGAACATTTGAACAGTACCCTCAAGCCACTCTCTCTCTAGTATCATCATCTCTC 1619

QY 549 GluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArg 568
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1620 CATATCGATGGTCTTATTTACTTCCCAACTCTAGCTACCTCGGAGATAGATTTGGAAGG 1679

QY 569 TrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrVal 588
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1680 TCGATCGATGAGGAGGAAGAACTGAGAAAGAGTGAAGAAACAGTCTCAATATACATC 1739

QY 589 IleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGlu 608
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1740 ATACTGTATTTGAGTGTGTGGGAACATPCGACGAGTGGTATGAATGCTCCAGGAA 1799

QY 609 ValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSerGlu 628
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1800 GTCAACAAATACTGGGACGAAGAGACCTCAAGCTAGTAGTAGCTAAACCCAGGAGCAGAG 1859

QY 629 ValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIleTyr 648
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1860 TTGATGAAGAAAGCTAAGCAAGTCCAAGTTCACAGAAACCATCGGCAAAAGACTGGATT 1919

QY 649 LeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerLys--Thr 667
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1920 CTCAGGTAGCAGAACCCGCTCTCAGCCTGTGATTTCGTGCTCCACACAGCTAAACCGGAC 1979

QY 668 AsnProLysLysAspGluThrGlu---GlyTrpAsnAsnVal 680
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1980 TCACCGGAGAAATCTCAGTGAACCGGAATTTTAAACAGTT 2021

RESULT 11
AR477457
LOCUS
DEFINITION
AR477457
VERSION
AR477457.1 GI:47234915
KEYWORDS
SOURCE
ORGANISM
REFERENCE
UnClassified.
1 (bases 1 to 1981)
AUTHORS
Allen,S.M., Falco,S.C. and Thorpe,C.J.
TITLE
Genes encoding sulfate assimilation proteins
JOURNAL
Patent: US 6696292-A 3 24-FEB-2004;
FEATURES
Location/Qualifiers
1..1981
source
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
```

Pred. No.:	2,39e-163	Length:	1981
Score:	2164.00	Matches:	403
Percent Similarity:	86.09%	Conservative:	86
Best Local Similarity:	70.95%	Mismatches:	79
Query Match:	62.27%	Indels:	0
DB:	6	Gaps:	0
US-10-762-049-18 (1-680) x AR477457 (1-1981)			
Qy	105	LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer	124
Db	8	GAGTCCGACCTGATCCCGCGCATCACCATGCCAGCTCGCCATCCCGCAGGGGCATCAGC	67
Qy	125	TyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIleProPro	144
Db	68	TACGCCAAGCTGCCAACCTGCGGCCCGTGCTCGGACTCTACTCGAGCTTCGTGCCGCCG	137
Qy	145	LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly	164
Db	128	CTGGTGTACCGCTGATGGGAGCTCCCAAGGACCTGGCGGGTGGGAGCGGTGGCGGTGGCG	187
Qy	165	SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeu	184
Db	188	TCGCTGCTCATCAGCTCCATGCTCGCGCAGCGAGGTGTCCGCCAGCGAGAACCCCGTGCTC	247
Qy	185	TyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGly	204
Db	248	TACTCTGACCTCGCCTTCAACGCCACCTTCTTCGCGCGGTCTTCACAGGCGCTCGCTCGGC	307
Qy	205	LeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMet	224
Db	308	CTCCTCAGGTGGCGCTTCATCGTGACCTGCTGTCGACGCGAGCATGGTGGGCTCATG	367
Qy	225	GlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis	244
Db	368	GC CGCGCGCGCAGCGTGGTGTGCTGCTGACGAGCTGAAGGGCATGTCTGGCGCTCGTCCAC	427
Qy	245	PheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlu	264
Db	428	TTCAACCACTTCCACCGACGTGCTCTCCGTCTAGGAATCCGCTTCACGCCAGACACACCG	487
Qy	265	TrpArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArg	284
Db	488	TGGCGGTGGAGAGCGTCTGCTCGGCTCGCGCTTCCTCTCTCTCTCTCTCTCTCTCTCTCT	547
Qy	285	TyrPheSerLysLysArgProArgPhePheTyrPheValSerAlaMetalProLeuThrSer	304
Db	548	TTTCATCAGCAAGAGGCGTCCCAAGCTGTCTGGATCTCCGCGGGCGCGCGCTTGACGTCC	607
Qy	305	ValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluVal	324
Db	608	GTCTGCTCTGGAGGCGTCTGCTGTACTCTACGACGCTGAAACACCAACCGCATCGAAGTG	667
Qy	325	IleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSer	344
Db	668	ATCGGTTACCTGNAGAAAGCCCTGAATCCACCGTCGGTGACAGCCCTGCAATCTCACCG	727
Qy	345	ProTyrMetThrThralaValLysThrGlyIleValValGlyIleIleSerLeuAlaGlu	364
Db	728	CCCTACATGATCTCGCGCTCAAGACTGGGATCATCACCGCGGTCTATGGCCCTCGCCGAA	787
Qy	365	GlyIleAlaValClyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLys	384
Db	788	GGAATCCCGTGGGAGGAGCTTCGCCATGTTCAAGAACTACCAATGACGGACACCAAG	847
Qy	385	GluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThr	404
Db	848	GAGATGATCGCATCGGACGATGAACGTCTCGGCTGCTCACGTCTGCTACCTGACC	907
Qy	405	ThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSer	424
Db	908	ACGGGCGCTTCTCGCGCTCCCGCTGAACATAACAACCGCGGTTCAGGACGGCATGTGCG	967

Qy	425	AsnIlelleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHis	444
Db	968	AAAGTGTGTCATGTCGTGGCGGTGATGGTCACGCTGCTGTCTCTGACCGCGCTGTTCAC	1027
Qy	445	TyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyIleuLeuAsp	464
Db	1028	TACACGCGCTGTGGTGTCTGGCGGATCATCGTCTCCGCGATGCTGGCGCCTGGTCGAC	1087
Qy	465	TyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSer	484
Db	1088	TTCCGGGCGCGGTGCACCTGTGGCGGTGCACAGGTGCACCTTTCGCTCTGGCGCGC	1147
Qy	485	AlaTyrIleGlyValValPheGlySerValGluIleLeuValIleValIleValIle	504
Db	1148	CGGTACCTGGCGCTCGTCTTCGGCAGCGTTCGAGGTTCGGCTCGCTCGCGCTCGCGCTC	1207
Qy	505	SerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIle	524
Db	1208	TCCTGTCTCGCGTCTCGTGTTCGTGCGCGCGCCAGGACCGGTCTCGGCAACATC	1267
Qy	525	ProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGly	544
Db	1268	CCCGGCACCATGTGTACCGAGATGGACAGTACGCCCGCGGACGCGTTCGCGCGC	1327
Qy	545	MetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGlu	564
Db	1328	GTGCTCGTGTGGCGTGCAGCGCGCCGCTACTTTCGCCAACGCCAGTACCTTCGCGAG	1387
Qy	565	ArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSer	584
Db	1388	AGGATCTCCGCGGTGCATCGACGAGGAGGCGCACCAAGAGCCAGGCGGAGATGGGC	1447
Qy	585	LeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSer	604
Db	1448	GTGCGGTACGTTGCTCTGCATCGGTGGTGCATCGGTAGATCGACAGCGCGGACGAGC	1507
Qy	605	MetLeuGluGluValLysLysIleThrGluArgGluLeuGlnLeuValLeuValAsn	624
Db	1508	ATGCTGGCAGAGTCAACAGTCTCTGGCAGGAGGGAATGCAGATCGTGTCTGGCGNAC	1567
Qy	625	ProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLys	644
Db	1568	CCGGCGCAGCAGATCATGAAGAAGCTGGACAGCTCCAAGGTGCTGGAGCAGATCGGCCAC	1627
Qy	645	LysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAla	664
Db	1628	GAGTGGGTGTTCGCGACGCTGGCGAGGCGGTGGCTGCGCGCATCTACGTGCTGCACTCG	1687
Qy	665	SerLysThrAsnProLysLysAsp	672
Db	1688	CACAGCGCGGAATGCCAAGGAC	1711
RESULT 12			
ATF4F15/c			
LOCUS	ATF4F15	123209 bp	DNA linear PLN 27-JAN-2000
DEFINITION	Arabidopsis thaliana DNA chromosome 3, BAC clone F4F15.		
ACCESSION	AL049711		
VERSION	AL049711.3	GI:6807566	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1		
AUTHORS	Alcaraz,J.P., Clabault,G., Cottet,A., Mache,R., Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F. and Salanoubat,M.		
JOURNAL	Unpublished		
REFERENCE	2	(bases 1 to 123209)	
AUTHORS	EU Arabidopsis sequencing, project.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-JAN-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried FRG, E-mail:		

lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
Coordinator: Marcel Salanoubat and Francis Quetier, Groupement
d'Interet Public, Centre National de Sequencage - GENOSCOPE; 2 rue
Gaston Cremieux, BP191, 91006 Evry Cedex, France;
http://www.genoscope.cns.fr
On Jan 28, 2000 this sequence version replaced gi:4691223.
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.

FEATURES
 source
 1..123209
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /variety="Columbia"
 /db_xref="taxon:3702"
 /chromosome="3"
 536..4710
 /gene="F4F15.10"
 complement(join(536..721,804..880,977..1062,1145..1335,
 1471..1757,1849..1975,2172..2288,2597..2710,2950..3015,
 3169..3340,4033..4240,4365..4710))
 /gene="F4F15.10"
 complement(join(536..721,804..880,977..1062,1145..1335,
 1471..1757,1849..1975,2172..2288,2597..2710,2950..3015,
 3169..3340,4033..4240,4365..4710))
 /gene="F4F15.10"
 /notes="Contains Prokaryotic membrane lipoprotein lipid
 attachment site AA369-379
 contains EST gb:T21459, Aa720146"
 /codon_start=1
 /product="sulfate transporter (AtST1)"
 /protein_id="CAB41310.1"
 /db_xref="GI:4678919"
 /db_xref="GOA:Q9SV13"
 /db_xref="UniProt/TREMBL:Q9SV13"
 /translation="MGEDYTFPGAEELHRRHTVEAPQPQPFLLKSLQYSVKETLFP
 DDPFFQKNQASRFGLGKFLPIFEWAPRYNLKFKKSDLIAGITITASLAIPQGIS
 YAKLANLPILGYSFVPLVYAVIGSSRDLAGVTAVASLLTGAMLSKEVDAEKDP
 KLYLHLPATFTFAGVLEASLGI FRLGFTIVDFLSHATIVGFMGGAATVVSLOQLKGI
 GLKHTDSDVLSVMSVFSOTHEWWSGVLCGCFLLSTLSTRVFSIKKPFVVA
 MAPLTSVLLGVFTHAERHGVQVIGDKGLNPLSGSDLIITSPYMSAVTKTGLI
 TGIIAIEGVAGRSFAMFNKNIIDGNKEMIAFGMNIIVGSFTSCYLLITGPPRSVAVN
 YNAGCTAMSNVMAIAVNTFLTLPLFHYPLVLSAIIISAMGLIDYQAAHLW
 KYDKDFLVCMSAYGVGVSVEIGLVAVASLARLLFVSRPKTAVKGNTPNSMIY
 RTEQVPSRTVPTGLILIDAPIYFANASYLRIRIWIDEEERVKSGSSLOYI
 ILDMSAVGNIDTSGISMMVEIKKVIDRRALKVLNPKGEVVKLTRSKFIGDLHGE
 WNFLTVEACSYMLHTFKTEPAKNEPNNV"
 complement(536..721)
 /gene="F4F15.10"
 complement(722..803)
 /gene="F4F15.10"
 /number=1
 complement(804..880)
 /gene="F4F15.10"
 /number=2
 complement(881..976)
 /gene="F4F15.10"
 /number=2
 complement(977..1062)
 /gene="F4F15.10"
 /number=3
 complement(1063..1144)
 /gene="F4F15.10"
 /number=3
 complement(1145..1335)
 /gene="F4F15.10"
 /number=4
 complement(1336..1470)
 /gene="F4F15.10"
 /number=4
 complement(1471..1757)
 /gene="F4F15.10"

gene
 536..4710
 /gene="F4F15.10"
 complement(536..721,804..880,977..1062,1145..1335,
 1471..1757,1849..1975,2172..2288,2597..2710,2950..3015,
 3169..3340,4033..4240,4365..4710)
 /gene="F4F15.10"

gene
 1471..1757,1849..1975,2172..2288,2597..2710,2950..3015,
 3169..3340,4033..4240,4365..4710
 /gene="F4F15.10"

CDS
 complement(join(536..721,804..880,977..1062,1145..1335,
 1471..1757,1849..1975,2172..2288,2597..2710,2950..3015,
 3169..3340,4033..4240,4365..4710))
 /gene="F4F15.10"

intron
 complement(1758..1848)
 /gene="F4F15.10"
 /number=5
 complement(1849..1975)
 /gene="F4F15.10"
 /number=6
 complement(1976..2171)
 /gene="F4F15.10"
 /number=6
 complement(2172..2288)
 /gene="F4F15.10"
 /number=7
 complement(2289..2596)
 /gene="F4F15.10"
 /number=7
 complement(2597..2710)
 /gene="F4F15.10"
 /number=8
 complement(2711..2949)
 /gene="F4F15.10"
 /number=8
 complement(2950..3015)
 /gene="F4F15.10"
 /number=9
 complement(3016..3168)
 /gene="F4F15.10"
 /number=9
 complement(3169..3340)
 /gene="F4F15.10"
 /number=10
 complement(3341..4032)
 /gene="F4F15.10"
 /number=10
 complement(4033..4240)
 /gene="F4F15.10"
 /number=11
 complement(4241..4364)
 /gene="F4F15.10"
 /number=11
 complement(4365..4710)
 /gene="F4F15.10"
 /number=12
 14327..15652
 /gene="F4F15.20"
 join(14327..14590,15098..15652)
 /gene="F4F15.20"
 /notes="similarity to heat shock transcription factor HSF30
 - Peruvian tomato, PIR2:S25480
 Contains HSF-type DNA-binding domain signature AA69-93
 contains EST gb:R65204, Aa042693"
 /codon_start=1
 /product="putative heat shock transcription factor"
 /protein_id="CAB41311.1"
 /db_xref="GI:4678920"
 /db_xref="GOA:Q9SV12"
 /db_xref="UniProt/TREMBL:Q9SV12"
 /translation="MMNPFLPEGCDPPPPPOPEGLHENAPPLTKTFEMVDDPNTD
 HIVSWNRGTSFVVDLHSHSTILLPRHFKSNFSSFIQLNTYGFRIEAEWRFAN
 BEFLGQQLLNKIKRNPFPTSSPSHDACNELRREKQVLMMEIVSURQOQTTKSY
 IKAMEQRLEGTERKORQWMSFLRAMQSPSLHOLLKQDKKIKLELDNESAKRKGS
 SSMSLEVLIALEMQGHGKQKNLBEEDHQLVVERELDDGFWEELLSDESLASTS"
 14327..14590
 /gene="F4F15.20"
 /number=1
 14591..15097
 /gene="F4F15.20"
 /number=1
 15098..15652
 /gene="F4F15.20"
 /number=2
 17211..18347
 /gene="F4F15.30"

```

gene      complement(join(17211..17351,17850..17927,18018..18178,
18272..18347))
/genes="F4P15.30"
CDS       complement(join(17211..17351,17850..17927,18018..18178,
18272..18347))
/genes="F4P15.30"
/notes="strong similarity to calmodulin - Tetrahymena
pyriformis (SGCS), PIRI:MCTE
contains EF-hand calcium-binding domain AA94-106;EF-hand
calcium-binding domain AA130-142
contains EST gb:R65197, R87041, Aa597913"
/codon_start=1
/product="putative calmodulin"
/protein_id="CA841312.1"
/db_xref="GI:4678921"
/db_xref="GOA:Q9S744"
/db_xref="UniProt/TrEMBL:Q9S744"
translations="MADAPTDEOIQFYBAFLCLDKSDGFIKELTKVMKSMGKNP
KASLOQMMSDVDFICNGGITTDDFLYIMAKNTQSASDELIEVRFDRDGDGLIS
QLELGGMKDMGKIKTAERAEHVMREADLDGDGFLSFHFESRMIAASY"
exon      complement(17211..17351)
/genes="F4P15.30"
/number=1
intron    complement(17352..17849)
/genes="F4P15.30"
exon      complement(17850..17927)
/genes="F4P15.30"
/number=2
intron    complement(17928..18017)
/genes="F4P15.30"
/number=2
exon      complement(18018..18178)
/genes="F4P15.30"
/number=3

Alignment Scores:
Pred. No.:      4,77e-150      Length:      123209
Score:          2026.50      Matches:      499
Percent Similarity: 41.40%      Conservative: 79
Best Local Similarity: 35.74%      Mismatches:   78
Query Match:    58.32%      Indels:       743
DB:             8           Gaps:         14

US-10-762-049-18 (1-680) x ATPF4F15 (1-123209)

QY      24  MetGlySerValAspTyrGluTyrProLeuGlyMetAsnPheGluArgValHisGln 43
Db      4710  ATGGGACGGAGGACTACACATTCCTCAGGAGCGGAGGAGCTTCACCGCCGCCACCAC 4651

QY      44  ---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db      4650  ACGGTGGAGGCTCCACAACCTCAACCGTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 4591

QY      63  ThrPhePheProAspAppProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db      4590  ACTCTGTGTTCCAGACGACCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 4531

QY      83  MetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLysTyrThrPheGln 102
Db      4530  GTGTAGGCTCAAACTACTCTCCCGATTTTCGATGGGCACACCGTACATCTCAAG 4471

QY      103  PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db      4470  TTCTTCAAAATCAGATCTCATCGCGGAATCACCATCGTAGTCTCGCCATCCCTCAGGGC 4411

QY      123  IleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeu-Tyr----- 138
Db      4410  ATCAGTTACGCCAAACTGCTAACTTGCCCCCAATTTCTGGCCCTTTGTAAAGCGTTATACA 4351

QY      138  ----- 138
Db      4350  TCCTTTATTTTATATATTTCCCTTAACAATTAAGGGTTTTTTGTGTGTATATACACT 4291

```

```

QY      139  -----SerSerPh 141
Db      4290  ATTTTTTGAAGTAGCGTTTGTACTTATGTAACTGTTTTTTTCGTTTAGATTTCGATTT 4231

QY      141  eIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVa 161
Db      4230  TGTACCGCATTTGGTATACGCGGTCTAGGAGTTCAAGGACTTACGGTGGAAACGGT 4171

QY      161  lAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValaAspProAsnGluAs 181
Db      4170  TCGGTTGCGTCTCTCTGTGACAGGTGCGATGCTGAGCAAAAGAGTTGATCTCGAGAAAG 4111

QY      181  pProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAl 201
Db      4110  TCTAAGCTTTACCTTCACCTTTCACCGCACCTTTTTTTCGCGCGGCTTCGGAAGC 4051

QY      201  aAlaLeuGlyLeuPhe----- 206
Db      4050  CTCCTTTGGAAT-TTTCAGGTTCTTTTACTTTATTTCTCCTTAATATTTTGTGCTTT 3992

QY      206  ----- 206
Db      3991  TTAATATGTTTTTACTTTATGATAAAAGTGTGAGTAATATATATTCGGTCCAATGATGGG 3932

QY      206  ----- 206
Db      3931  AAAGATTCCGGCATAAAAAATATCAACCAATAGTAACTTTTTCTTACCTTGATTCTT 3872

QY      206  ----- 206
Db      3871  AAGATCTAAAAATTCATAAGCACTATGTACTATGTACTATATGAGAGTGAATTCGTCAT 3812

QY      206  ----- 206
Db      3811  CGCGAGTTACGTCACTGAGAAACAAGATACAAGAGCTGAGACAAACAACAATAGGACAA 3752

QY      206  ----- 206
Db      3751  AGCTAGATAACTATTAGAAAAATCACACCGTGGTATGCTCGGTTTAATAATACATATATAC 3692

QY      206  ----- 206
Db      3691  ATGTAACAGTGGAGAAATCATACTAGAAATATGTAATTACTTTACTCGTAAGTTGGGTATA 3632

QY      206  ----- 206
Db      3631  ATTTAATACATATCTGCCCAAGTTTATTTTGTATTCCTCGTATCATATCATAGTGTGATT 3572

QY      206  ----- 206
Db      3571  GATTCAAAAGCACAGAGAGATATATATATATTATAGGTAGAAATCTTTTATTTTGGTTC 3512

QY      206  ----- 206
Db      3511  CACAAAATTTTCGTATATCTAGCCAGAACTGGATTGAAGCATTAAGGTATTAACTAAGTT 3452

QY      206  ----- 206
Db      3451  GGTGAGTTTGATGATGATGTGAGTTTAATAGTGCAAACTGATAAATGATTGTTTAGTAG 3392

QY      207  -----ArgLeuGlyLe 210
Db      3391  CTTTAAAAATAAAATTAAGAGAAGAAAGAAATGATGAAAAGGGAATTCAGAGTTAGGGTT 3332

QY      210  urleValaAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrVa 230
Db      3331  CATAGTGGACTTTCTATCTTCATCGCAACGATAGTAGGATTCATGGGAGGAGCAGCAGGT 3272

QY      230  lValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAs 250
Db      3271  GGTGAGTCTGCAACAGCTTAAGGGTATTTTCGGACTTTAAACATTTTCACAGACTCTACCGA 3212

```



```

/db_xref="GI:3068706"
/translation="MAARLLFRSSQILRSIQNRPNQISSFSPSPCPFIHSLTTSAPD
PSRLSLTLRSLSIARRGTPRKIDIGAKAROMNRRLWTYALTFSCIAGFVVIVL
NOFQDOLVFLTPSDAMEKFAENPTKNKFRIGLGLVLEGSVAQPAASQEMEFVITDLIT
DILVRKGSISLPDLFRGHSVVVGEFIKPTDEVKRVSTKPISKARNLDCFSATEV
LAKHDKYMPQEVAAAEIKKKIIEAARATEQAEEVAAS"
complement (13443..16043)
/gene="AtG2"
complement (join(13443..13716,13805..13978,14067..14158,
14245..14381,14467..14682,15169..15322,15415..15564,
15660..15744,15854..16043))
/gene="AtG2"
/product="putative nuclear DNA-binding protein G2p"
complement (join(13640..13716,13805..13978,14067..14158,
14245..14381,14467..14682,15169..15322,15415..15564,
15660..15744,15854..15947))
/gene="AtG2"
/note="similar to Mus musculus p38-2G4 protein; SwissProt
Accession Number P50580"
/codon_start=1
/product="putative nuclear DNA-binding protein G2p"
/protein_id="AAC14407.1"
/db_xref="GI:3068707"
/translation="MSSDERDEKELSLTSPVVVTKYKAAEIVNKALQVLAECKPK
AKIVDICEGDSFIKEQTASMYKNKKKIERGVAFPTCISVNNVTHGFSPLASDSVL
EPDMMVKIDMGCHIDGFIALGVHTVLQEPGLSGKADYIAAANTAADVALRLVRPGK
KNTVDTEATQKVAADYCKIVEGLVSHQKHQVIDGNKVVLVSANSPETTVDBVEFEEN
EYVVIDVASTGDKPKLDEKOTTYIKKDESNNYOLNKKASRFIISIKQNPMPF
TARLSEKPARLGVECVNHHLOLPYPVLVYKPGFVAQIKETVILLMNGSDRIISHT
LQELPKTITDEPFIKGLWALGIIKKKGGKKKAQKAGKEASTEAEPEMDASSNAQE
"
17135..18031
/gene="ATEM1"
join(17135..17314,17500..18031)
/gene="ATEM1"
/product="Em1 protein"
join(17197..17314,17500..17840)
/gene="ATEM1"
/codon_start=1
/product="Em1 protein"
/protein_id="AAC14408.1"
/db_xref="GI:3068708"
/translation="WASKQLSREELDEKAKQGETVVPQGTGGHSLAQVHLAEGRSG
GQTRKEQLGHEGYQETIGHKSGEQLGHEGYQEMHGKGGEARKEQLGHEGYQEMGH
KGGEARKEQLGHEGYQEMGRKGGLSTMEKSGGERABEEGIEIDESKFTNK"
complement (18488..21370)
/gene="AtG4"
complement (join(18488..18742,18838..18885,18970..19046,
19410..19554,19644..19719,19798..19859,19956..20039,
20114..20179,20253..20339,20428..20513,20596..20683,
20868..20953,21037..21081,21163..21370))
/gene="AtG4"
complement (join(18680..18742,18838..18885,18970..19046,
19410..19554,19644..19719,19798..19859,19956..20039,
20114..20179,20253..20339,20428..20513,20596..20683,
20868..20953,21037..21081,21163..21313))
/gene="AtG4"
/note="similar to Rhodobacter capsulatus
bacteriochlorophyll synthetase BchG encoded by GenBank
Accession Number Z11165"
/codon_start=1
/product="putative chlorophyll synthetase"
/protein_id="AAC14409.1"
/db_xref="GI:3068709"
/translation="WTSILNTVSTIHSRSVTSVDVGVLSLRNSDSVEPTRRRSGFST
LIESGFRVVRARIEDTDKVSQTPDKAPAGGSINOLLGIGKASQETNRKWLRLQ
LTKPVTPPLVWGVCGAASNFHWTPEDVAKSILCMMSGPCUTGYTGTINDWYDR
DIDANETPRSPGASISRPVITQWVLLGLAGTAGLIDVAGHTPTTFVYALGG
SLSLXYTSPKSLQKNGWGVNFALCASISLPWAGQALFGLTLPDVPVLLYSIA
GLGIATVNDPKFVEGDLQSLPVARFQETAKVICVCAIDITOLSVAGYLLASGP
YVALALVALIIPQIVFQPKYFLKDPVKYDKVQAQAQPLVLIGIFVITALASQH"
22568..27625
/gene="AtG5"

```

mRNA

```

join(22568..22766,22902..22976,23067..23194,23297..23388,
23469..23526,23639..23695,23775..23808,23947..24027,
24107..24213,24563..24649,24731..24808,24982..25046,
25268..25366,25516..25615,25979..26087,26196..26268,
26601..26694,26792..26875,27183..27249,27344..27625)
/gene="AtG5"
/product="putative transmembrane protein G5p"
join(22616..22766,22902..22976,23067..23194,23297..23388,
23469..23526,23639..23695,23775..23808,23947..24027,
24107..24213,24563..24649,24731..24808,24982..25046,
25268..25366,25516..25615,25979..26087,26196..26268,
26601..26694,26792..26875,27183..27249,27344..27471)
/gene="AtG5"
/note="similar to yeast Sac1 protein"
/codon_start=1
/product="putative transmembrane protein G5p"
/protein_id="AAC14410.1"

```

CDS

```

join(22616..22766,22902..22976,23067..23194,23297..23388,
23469..23526,23639..23695,23775..23808,23947..24027,
24107..24213,24563..24649,24731..24808,24982..25046,
25268..25366,25516..25615,25979..26087,26196..26268,
26601..26694,26792..26875,27183..27249,27344..27471)
/gene="AtG5"
/note="similar to yeast Sac1 protein"
/codon_start=1
/product="putative transmembrane protein G5p"
/protein_id="AAC14410.1"

```

Alignment Scores:

```

Pred. No.: 5,7e-150 Length: 60000
Score: 2020.50 Matches: 497
Percent Similarity: 41.40% Conservative: 81
Best Local Similarity: 35.60% Mismatches: 78
Query Match: 58.14% Indels: 743
DB: 8 Gaps: 14

```

US-10-762-049-18 (1-680) x AF049236 (1-60000)

```

QY 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
DB 58057 ATGGGACGCGAGCATACACATTCCCTCAAGGAGCGGAGGAGCTTCCACGCGCCACAC 57998
QY 44 ---ValGluValProProGlnProPhePhePheLeuLysSerLeuLysTyrSerLeuLysGlu 62
DB 57997 ACGGTGGAGGCTCCACACCTCAACGTTCTTGAAGTCACTTCAGTACTCAGTGAAGAA 57938
QY 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
DB 57937 ACTCTGTTTCCAGACGACCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 57878
QY 83 MetLeuGlyLeuGlnPhePhePheProIlePheGluTyrAlaProIlyTyrThrPheGln 102
DB 57877 GTGTTAGCGCTCAAAATACTTCTCCCGATTTTCGAATGGGACCGCTACCAATCTCAAG 57818
QY 103 PheLeuLysAlaAspLeuLeuAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
DB 57817 TTCCTCAATCAGATCTCATCGCCGGAATCACCATCGTAGTCTGCCATCCCTCAGGCG 57758
QY 123 IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeu-Tyr----- 138
DB 57757 ATCAGTTAGCCCAACTTGCTAACTTGCCCCCAATTTCTTGGCCTTTGTAAAGCGTTATACA 57698
QY 138 ----- 138
DB 57697 TCTTTTATTTTATTTATTTTCCTTAACAATTAACGGGTTTTTTTGTGTGTATATACACT 57638
QY 139 -----SerSerPh 141
DB 57637 ATTTTTCAGTAGCGTTTGTACTTATTGTAACTGTTTTTTTGTGCTTTAGATTCAGATT 57578
QY 141 eileProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVa 161
DB 57577 TGTACCGCATTTGGTATACGCGGTGCTAGGGAGTTCAAGGAGCTTAGCGGTGGAAACGCT 57518
QY 161 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAs 181
DB 57517 TCGCGTTCGCTCTGTTGACAGAGTGCATGCTAGCAAAAGAAAGTGTGCTGAGAAAGA 57458
QY 181 pProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAl 201
DB 57457 TCCTAAGCTTTACCTTCACCTTGCTTTTCCGCCCACTTTTTCGCGCGGCTTCTCGAAGC 57398
QY 201 aAlaLeuGlyLeuPhe----- 206

```


Db	57397	[::: :::]	57338
Qy	206	CTCTCTGGAAATTTTCAGGTTCTTTTTTACTTTTATTTCTCCTTAATTAATTTTTTGTCCTTT	288
Db	57337	TAATATGTTTACTTATGATAAAAGTGTGAGTAATATATATTCGGTCCAATGCATGTGGGA	56199
Qy	206	-----	288
Db	57277	AAAGATTCCGGCATAAAAAATATCAACAAATAGTAACCTTTTTCCTTACCTTGATTTCTTA	56139
Qy	206	-----	288
Db	57217	AGATCTAAAAATTCATAAGCACTATGTACTATGTACTATATGAGAGTGAATCGGTCATC	56079
Qy	206	-----	300
Db	57157	GCGAGTTACGTCAGTGAACAAAGATACAAAGAGCTGAGACAACAAACCAATAGGACAAA	56019
Qy	206	-----	320
Db	57097	GCTAGATAACTATTAGAAAAATCACACCGTGGTATGCTCGGTTTAAATACATATATACTACA	55959
Qy	206	-----	325
Db	57037	TGTAACAGTGGAGAAATCATACTAGAAATAGTAATTACTTTACTTGTGAAGTGGGTATAA	55899
Qy	206	-----	325
Db	56977	TTTAATACATATCTGCCAAGTTTATTTTGTATCCCGTATCATATAATCATAGTGTGATGA	55839
Qy	206	-----	325
Db	56917	TTCAAAAGCACAAGAGAGATATATAGTATATATATAGTGAATCTTTTATTTTGGTTCCA	55779
Qy	206	-----	325
Db	56857	CAAAATTCGGTATATCTAGCCAGAACTGGATTGAAGCATAAAGGTATTAACTAAGTTGG	55719
Qy	206	-----	325
Db	56797	TGAGTTTGATGATGTGAGTTTATATAGTGCAAACTGTAAATTGATTTGTTTAGTAGCT	55659
Qy	207	-----	337
Db	56737	TTAAAAATAAAATTAAGAGAAGAAAGAAATGATGMAAAGGGAATTCAGGTTAGGGTTCA	55600
Qy	211	leValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValV	357
Db	56677	TAGTGGACTTTCTATCGCATGCAACGATAGTAGGATTCATGGAGGAGCAGCCGCTGG	55540
Qy	231	alCysLeuGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAspI	363
Db	56617	TGAGTCTGCAACAGCTTAAGGTATATTCGGACTTAAACATTTTCACAGACTCTACCGATG	55481
Qy	251	leIleSerValMetArgSerValPheThrGlnThrHisGlu	363
Db	56557	TTATCTGTGTCATGCGTTCGGTTTCTCTCCAAAACCTCACGAGGTGCTCTCATTTCCGGTTA	55421
Qy	264	-----	363
Db	56497	TGTTGCTAATACATCCAAAATATCATTTTCAATGTGATGATTTGTTCTCTCAATTAATAC	55361
Qy	264	-----	371
Db	56437	TATGAATATTCATAGTATTATTAACAGTAAACGCTCTAAACATGTGTTTGTGTTGTTTG	55301
Qy	265	-----	391
Db	56377	TTTTTGTGTAGTAGAGATGGAAAGTGGCGTCTCTTGGATGTGGTTTCTATCTTCTTC	55241
Qy	280	eulSerThrArgTyrPheSerLys-----	404
			55181

```
QY 404 ----- 404
Db 55180 CACATAGTCACATATGGTTATTAACTATGAAAGAAAGCTTATATATTAAATTACGACG 55121
QY 405 ----- Thr-GlyProPheSerArgSerAlaValAsnTyrAenAlaGlyCysLys 420
Db 55120 TACCATACTATACGAGACCATTTTCAAGGTCGGCAGTGAACATACAAACGGGGTTGCAAG 55061
QY 421 ThrAlaAsnSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db 55060 ACCGCAATGTCCAACATAGTAGGCGATGCGGGTATGTCACATCTCTTCCTCCAC 55001
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db 55000 CCGCTTTTTCACATACACACCATCGTCCTCTCTGCCATCATATCCGCAATGCTC 54941
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 54940 GGACTCATTTGACTATCAAGCTGCCATCCATCTCTGGAAAGTTGACAAAGTTCCGACTTCTC 54881
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db 54880 GTCTGATGAGCCCTACGTGTGGGTGCGTATTCGGCAGTGTAGAGATTGACTGTCGTA 54821
QY 501 Ala----- 501
Db 54820 GCGGTTAGTATATACTTTATATGCTGTAATGAGCCTAGATGCCATGCTCTTTTTC 54761
QY 501 ----- 501
Db 54760 CTTCTATTTGCGTTGGCATTTTAAACATTCGCAATATTATTAATTTCTAAATGTTTTT 54701
QY 502 -----lleValIleSerValLeuArgValLeuLeuPheIleAlaArg 515
Db 54700 TTTTTCATGTTGATGAAGTGGCGATATCTATAGCAGGTTGTGCTGTGTGTCGAGG 54641
QY 516 ProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHis 535
Db 54640 CCAAAACTGCGGTGAAGGAAACATACCAACAGCATGATCTATAGGAACACTGAGCAG 54581
QY 536 TyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyr 555
Db 54580 TACCCATCATCAAGAACCGTCTGTTGATTTCTCATCTTGAGATTGATGCTCCCATCTAC 54521
QY 556 PheAlaAsnAlaSerTyrLeuArgGlu----- 564
Db 54520 TTTGCTTAATGCCAGTTACTTGCGTGAAGTAAGTAACCTAACCCCTCATTCGTATTACATGCC 54461
QY 565 -----ArgIleThrAr 568
Db 54460 AAAGATTGAAGATGTTTTTGAATAATGTTTACTTTTCAATAAATCTACAGATCATAG 54401
QY 568 gTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrVa 588
Db 54400 GTGGATTGATGAGAGAAAGAGAGAGTTAAACAATCAGGAGAGAGCAGCTTACAATATAT 54341
QY 588 lIleIleAspMet----- 592
Db 54340 TATACTCGATATGTCAGGTAATGCATATGATCTTGATTTTCTTAACACTAGCTAATGTT 54281
QY 593 -----SerAlaValG1 596
Db 54280 GTTTTCGTCGAGCGCTTACTTAGTTACTCTGTGCGGTTTGTGTTAAACAATCAGCTGTTGG 54221
QY 596 yAsnIleAspThrSerGlyIleSerMetLeuGluValLysLysIleThrGluArgAr 616
Db 54220 TAATATCGACAAAGCGGTATTAGCATGATGTTGGTGAATTTAAGAAAGTCAATTCACAGGAG 54161
QY 616 gGluLeu----- 618
Db 54160 AGCGTTAAAGGTAAACGTAGTGGGTGATTAATATACATTTTATAGGATATGTAACGCGTTTCTT 54101
```

```
QY 619 -----GlnLeuValLeuValAsnProValSerGluV 629
Db 54100 CTTGTTAAATATCTTTCTTTTAAATCCACAGTTGGTATTGTCAAATCCAAAGGAGAGG 54041
QY 629 alMetLysLysLeuAsnLysSerLysPhe---GlnAsnHisLeuGlyLysLysTyrIleT 648
Db 54040 TCGTGAAGAAATTAACACGATCCAAATTCATCGGTGATCATTTGGGCAAGAGTGGATGT 53981
QY 648 yrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerLysThrA 668
Db 53980 TCTTAACGCTAGGAAGACGAGCGGTGTAGCTACATGCTTCACACGTTTAAACCG 53921
QY 668 snProLysLysAspGluThrGluGlyTyrAsnAsnVal 680
Db 53920 AACCG---GCCCTCCAAACAGACGCTTGGAAACACGTA 53886
RESULT 14
AB012048 7012 bp DNA linear PLN 14-FEB-2004
LOCUS Arabidopsis thaliana gene for sulfate transporter, complete cds,
DEFINITION clone:AST12.
ACCESSION AB012048
VERSION AB012048.1 GI:2967455
KEYWORDS sulfate transporter.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1
AUTHORS Takahashi,H., Sasakura,N., Kimura,A., Watanabe,A. and Saito,K.
TITLE Identification of two leaf-specific sulfate transporters in
Arabidopsis thaliana
JOURNAL Plant Physiol. 121, 686 (1999)
REFERENCE 2 (bases 1 to 7012)
AUTHORS Arabidopsis thaliana
TITLE Direct Submission
AUTHORS Takahashi,H., Sasakura,N. and Saito,K.
JOURNAL Submitted (06-MAR-1998) Hideki Takahashi, Chiba University, Faculty
of Pharmaceutical Sciences; 1-33, Yayoi-cho, Inage-ku, Chiba, Chiba
263-8522, Japan [E-mail:htaka@p.chiba-u.ac.jp, Tel.81-43-290-2906,
Fax:81-43-290-2905]
FEATURES
source
location/Qualifiers
1..7012
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="AST12"
/ecotype="Columbia"
join(1510..1855,1980..2187,2880..3051,3205..3270,
3510..3623,3973..4053,4250..4376,4468..4754,4890..5080,
5163..5248,5346..5422,5505..5690)
/codon_start=1
/product="sulfate transporter"
/protein_id="BAA25175.1"
/db_xref="GI:2967456"
/translation="MGTEDTYFPQGAELHRRHHTVEAPQPOPPLKLSQYSVKETLFP
DDPPEFQNONAKFVLGLKYLFIPEWAPRYNLKFKFSKLSLGIAGITIASLAIPGIGS
YAKLANLPILGLKYSFPPLVYAVLGSRDLAVGTVAVASLLTGAMLSKEVDASKDP
KLYLHLATATFFAGVLEASIGIFRLGFIVDLFSHATIVGFMGGAATVVSUQLKGIF
GLKHFDSTDIVISVMRSYFSTHERWESGLVGCGLFLLSTRYFSIKKPFVVA
MAPLSTVLGSLVYFTHAERHGVQVSDLIFTSPYMTAVKTLGITIGIIIAELGVAV
GRSPAFKNNIIDGNKEMIAFQNNIVGSFTSCYLTGTFPSRSVAVNYAGCKTAMSN
VMAVMTFLPLPLFHYTLPLVLSAIIISAMGLIDYQAIHLWKVDKDFPLVCM
AYGVVFSVIGLVVAISIRALLLFVSRPKTAVKGNINSMIYRNTVEQYPSRTV
PGILILEDAPYFANASYLRIRWIDEERVKQSGSLQYIILDSMAGNIDT
SGISNMVIRKVIDRRALKVLNSNPKGEVVKLLTRSKFIGDHLGKEMWFLTVGEAVEA
CSYMLHTFKTEPASKNEPWNV"
ORIGIN
```

Alignment Scores:

Pred. No.:	1.57e-150	Length:	7012
Score:	2012.50	Matches:	500

Percent Similarity:	41.39%	Conservative:	79
Best Local Similarity:	35.74%	Mismatches:	77
Query Match:	57.91%	Indels:	745
DB:	8	Gaps:	14
US-10-762-049-18 (1-680) x AB012048 (1-7012)			
QY	24	MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln	43
Db	1510	ATGGGCACGGAGGACTACACATTTCCCTCAAGGACGGAGGAGCTTCACCCGCCACCCAC	1569
QY	44	---ValGluValProProGlnProPhePhePheLeuLysTyrSerLeuLysGlu	62
Db	1570	ACGTGGAGGCTCACAACTCAACCTTCTTGAAGTCATCTCAGTACTCAGTGAAGAA	1629
QY	63	ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysPhe	82
Db	1630	ACTCTGTTCCAGACGACCTTTTAGACAATTTAAGAACCAAAATGCATCAAGAAATTT	1689
QY	83	MetLeuGlyLeuGlnPhePhePheProIlePheGluTyrAlaProLysTyrThrPheGln	102
Db	1690	GTGTAGGCTCAAACTTCTCCCGATTTTCGAATGGCACACCGCTACAACTCTCAAG	1749
QY	103	PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly	122
Db	1750	TTCTTCAATTCAGATCTCATGCGGGAATCACCATCGCTAGTCTCGCCATCCCTCAGGGC	1809
QY	123	IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeu-Tyr	138
Db	1810	ATCAGTTACGCCAACTTGCTACTTGCCCCCAATCTTGCCCTTGTGAAGCGTTATACA	1869
QY	138	-----	138
Db	1870	TCCTTTATTTTATTAFTTTTCTTAACAATTAACGGGTTTTTTGTGTGTATATACACT	1929
QY	139	-----SerSerPh	141
Db	1930	ATTTTGTGAAGTAGGTTTGCTACTTATGTAACTGTTTTTTTTCGGTTTAGATTTCGAGTTT	1989
QY	141	eIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVa	161
Db	1990	TGTACCCGCAATGGTATACGCGGTGTAGGAGTTCAAGGGACTTAGCGGTGGAAACGGT	2049
QY	161	lAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAs	181
Db	2050	TGCGGTTGCGCTCTGTTCAGAGGTGCGATGCTCAGCAAAAGAGTTGATGCTGAGAAGA	2109
QY	181	pProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAl	201
Db	2110	TCCTAAGCTTTACCTTCACCTTGCTTCCCGCCACTTTTTTCGCGCGGCTTCTCGAAGC	2169
QY	201	aAlaLeuGlyLeuPhe	206
Db	2170	CTCTCTGGAAAT-TTTCAGGTTCTTTTTCATTATTTCTCCTTAATATTATTTGTCCCTT	2228
QY	206	-----	206
Db	2229	TTAATATGGTTTTACTTATGATAAAAGTGTGAGTAATATATTCGGTCCAATGATGTGG	2288
QY	206	-----	206
Db	2289	AAAAGATTCCGGCATAAAAAATATCAACCAATAGTAACTTTTTTCCTTACCTTGATTTCTT	2348
QY	206	-----	206
Db	2349	AAGATCTAAAAATTCATAGCACTATGTACTATGTACTATATGAGAGTGAATTCGTCAT	2408
QY	206	-----	206
Db	2409	CGCGAGTTACGTCACGTGAAAACAAGATACAAGAGCTGAGACAACAAACCAATAGGACAA	2468
QY	206	-----	206
Db	2469	AGCTAGATAACTATTAGAAAAATCACACCGTGGTATGCTCGGTTTAAATACATATATACTAC	2528
QY	206	-----	206
Db	2529	ATGTAACAGTGGAGAAATCATAGAAATAATGTAAATTACTTTACTGTGAAGTTGGGTATA	2588
QY	206	-----	206
Db	2589	ATTTAATACATATCTGGCCAAAGTTTATTTTGATTTCCCGTATCATATAATCATAGTGTGATT	2648
QY	206	-----	206
Db	2649	GAITCAAAAGCACAGAGAGATATAGTATATTATAGGTAGTATCTTTTATTTTGGTTC	2708
QY	206	-----	206
Db	2709	CACAAAATTTTCGGTATATCTAGCCAGAACTGGATTGAAGCATAAAGGTATTAACTAAGTT	2768
QY	206	-----	206
Db	2769	GGTGAAGTTTGATGATGATGTGAGTTTAAATAGTGCAAACTGATAAAATTTGATTTAGTAG	2828
QY	207	-----ArgLeuGlyLe	210
Db	2829	CTTTAAAAATAAAATTAAGAGAAGAAAGAAATGATGAAAAGGAATTCAGGTTAGGTT	2888
QY	210	uIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyValalaThrVa	230
Db	2889	CATAGTGGACTTCTATCGCATGCAACGATAGTAGGATTCATGGGAGGAGCAGCGGT	2948
QY	230	lValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAs	250
Db	2949	GGTGAAGTGCACACAGCTTAAGGATTTTTCGAGCTTAAACATTTTCACAGACTCTACCGA	3008
QY	250	pIleIleSerValMetArgSerValPheThrGlnThrHisGlu	264
Db	3009	TGTTATCTCTGTCATGCGTTCCGTTTTTCCCAAACTCAGAGGTGTCCTCTCATTTCCGT	3068
QY	264	-----	264
Db	3069	TATGTTGCTAATACATCCAAAATATCATTTTCAATTCGATGATGTTCTCTCTTAATAAT	3128
QY	264	-----	264
Db	3129	ACTATGAATATTTCATAGTATTATTAAACAGTAAACGCTCTAAAACATGTGTTTGTGTT	3188
QY	265	-----TyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePh	279
Db	3189	TGTTTTTGTGTTGTAGTGGAGATGGGAAAGTGGCGTTCTTGGATGGGTTTTCTATTCTT	3248
QY	279	eLeuLeuSerThrArgTyrPheSerLys	288
Db	3249	TCCTCTCCACCAGATAITTCGT-AAAGATTATAATTTGTTATACCTTTTATATATCAATTTA	3307
QY	288	-----	288
Db	3308	AATAAGTGAATAATAACACATTTTATTTCGAATAGCAGCTACAATGATCTTTTCGAATA	3367
QY	288	-----	288
Db	3368	ATTCCATTTACTTTTACTCTCTCGCTTTTTCATCTTTCTAGAAAAGATAAAAAAGTAT	3427
QY	288	-----	288
Db	3428	ACGATATGGTGGACTGAAAATGTAAATAACTACACAATCAAAACGTTATTAAATGTGTTT	3487
QY	289	-----LysArgProArgPhePheTyrPheValSerAlaMe	299
Db	3488	AATGAATGAATGATGATGAGAGCATCAAGAAACCAAAATTTCTTTTGGGTGGCGCAT	3547
QY	299	tAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGlyL	319
Db	3548	GGCTCCTTTGACCTCAGTGAATTTCTTTGGAAAGTCTCTTTGGTTTACTTCTCACTCAGCTGAGAG	3607

```
QY 319 sHisGlyValGluVal----- 324
Db 3608 ACATGGTGTCAAGTGGTATTACTCAATTTCTAGTCTAGNAATCTCCTTATCTGATGATAA 3667
QY 324 ----- 324
Db 3668 ATAATCAAAATAAATAATTTTCAAGTCAAGACCAAGAGATAACGTTGATGATTTTAC 3727
QY 324 ----- 324
Db 3728 CGAATCTTAAGACATTATATATTTTAAAGAAAAGAAAAAGTTAGGTAGAATATTAT 3787
QY 324 ----- 324
Db 3788 GGTAAGAAAAAGAACAAAGACATGAAGTTGAAGTAGTAAGGAATTTGGRAAAAAGAG 3847
QY 324 ----- 324
Db 3848 TGGGCTCATGGGCTTTTGTGTAGTTTGTTCGAAAGGGAGAGGAATTTATCGAATAAAAGGAA 3907
QY 325 -----IleGlyGluLeuLeuLysGlyLeuAsnProp 335
Db 3908 ATTAAGGTATACGATTGTGATGTGCAGATAGGGACCTGAAAGAAAGGGTTGAATCCAC 3967
QY 335 roSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyI 355
Db 3968 TCTCCGGTTCGTATCTCACTTTACTTCCCTTACATGTCCACAGCTGTCAAAACTGGCC 4027
QY 355 leValValGlyIleIleSerLeuAla----- 363
Db 4028 TCATCACTGGCATCAATGTCTCGCTGTAAAGCCCATTTATCCCTTTTCTTATTATTAGGCTT 4087
QY 363 ----- 363
Db 4088 TGATAGTTATCTTATTATTACCAAGTGAATTTAAACAAATTAAGGCTTTATACCTA 4147
QY 363 ----- 363
Db 4148 AGGCTTAAATCACCACACACATATGGGAATATGATTTTCAATGGTGAAGAGTGAAGAC 4207
QY 364 -----GluGlyIleAlaValGly 369
Db 4208 AAAAACTCATCAATCAAAATCAAAATTCATGTGTAAAAATGTAGGAAGAGTAGCAGTGGGG 4267
QY 370 ArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIle 389
Db 4268 AGAGATTTTGCATGTTCAGAACTACAACATAGACGGGAACAAGAGATGATAGCGTTT 4327
QY 390 GlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThr----- 404
Db 4328 GGAATGATGAACATCGTTGGTTCCTTCACATCTTGTATTACCTCACACCCGGTACCTACACA 4387
QY 404 ----- 404
Db 4388 CTCACACATAGTCACATATGGTTATTAACTATGAAAGAAAAAGCTTTATATATTATTAC 4447
QY 405 -----Thr-GlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCy 419
Db 4448 GACGTACCATACTATACAGACCATTTTCAAGTCGGCAGTGAACACTACAACCGCGGTTG 4507
QY 419 sLysThrAlaAlaSerAsnIleMetSerLeuAlaValMetLeuThrLeuLeuPheLe 439
Db 4508 CAAGACCGCAATGTCCAACATAGTATGGCGATTGCGGTATTGTTCCACTCTCTCTCT 4567
QY 439 uThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMe 459
Db 4568 CACACCGCTTTTTCATACACACACTCGTCGTCCTCTCTGCGCATCATCATATATATAT 4627
QY 459 tLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLeuValAspLysPheAspPh 479
Db 4628 GCTCGGACTATGATATCAAGTCGCCATCCATCTCTCGAAAGTTGACAAAGTTTCGACTT 4687
```

```
QY 479 eValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVa 499
Db 4688 CCTCGTCTGCATGAGCGCTACGTTGGGTCGTATTCGCGAGTGTAGAGATTGGACTCGT 4747
QY 499 lIleAla----- 501
Db 4748 CGTAGCGGTAGTATATATCTTTATAATGCTGAATGAGCCTAGATGCCACTGCTCTTTT 4807
QY 501 ----- 501
Db 4808 TTTCCTCTTATTGCGTTTGGCAATTTAACCTTGCATATTATTAAATTTCAAATGTT 4867
QY 502 -----IleValIleSerValLeuArgValLeuPheIleAl 514
Db 4868 TTTTITTTTTCATGGTGTAGAGGTCGCGATATCTATAGCGAGGTTGCTGCTGTTGTGTC 4927
QY 514 aArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValG 534
Db 4928 GAGGCCAAAACCTCGCGTGAAGGAAACATACCAACAGCATGATCTATAGGAACACTGA 4987
QY 534 uHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaProI 554
Db 4988 GCAGTACCATCATCAAGAACCGTTCTCTGTATTCTCATCTTGGAGATTGATGCTCCCAT 5047
QY 554 eTyrPheAlaAsnAlaSerTyrLeuArgGlu----- 564
Db 5048 CTACTTTGCTAATGCCAGTTACTTCGCTGAAAGGTAACCTAACCCCTCATTCGTTATTACA 5107
QY 565 -----ArgIleT 567
Db 5108 TGCCAAAGATTGAAGAATGTTTGAATAATGTTTACTTTTCAATAAATCTACAGAATCA 5167
QY 567 hrArgTTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnT 587
Db 5168 TAAGTGTGATTGATGAAGAGGAAGAGAGTGTAAACAATCAGAGAGAGCAGCTTACAA 5227
QY 587 yrValIleIleAspMet----- 592
Db 5228 ATATTATATCTGATATGTCAGGTAATGCATATGATGATCTTGTATTCTCTAACACTAGCTAA 5287
QY 593 -----SerAla 594
Db 5288 TGTGCTTTTTCGTTGAGGCTTACTTAGTTACTCTCTGCGTGTGTTGTTTAAACAAATCAG 5347
QY 595 ValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGlu 614
Db 5348 GTTGTATATATCGACAAAGCGTATTAGCATGATGTTGGTGAATTAAGAAAGTCATTGAC 5407
QY 615 ArgArgGluLeu----- 618
Db 5408 AGGAGAGCGTTTAAAGGTAACGTAGTCGCTGTGATAATAATACATTTTATAGGATATGTAACGCGT 5467
QY 619 -----GlnLeuValLeuValAsnProValSe 627
Db 5468 TTCTTCTTGTATAATATTCTTCTTTTAAATCCACAGTTGGTATTGTCCAAATCCAAAGG 5527
QY 627 rGluValMetLysLysLeuAsnLysSerLysPhe--GlnAsnHisLeuGlyLysLysTr 646
Db 5528 AGAGTGTGGAAGAAATTAACAGATCCAAATTCATCGGTGATCATTTGGGCAAGAGTGT 5587
QY 646 pIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerLys 666
Db 5588 GATGTTCTTAACCGTAGGAGAGCAGTGGAGGCTTGTAGCTACATGCTTCACACGTTTAA 5647
QY 666 sThrAsnProLysLysAspGluThrGluGlyTTrpAsnAsnVal 680
Db 5648 AACCGAACCG---GCCTCCAAAACGAGCCTTGGAAACACGTA 5687
RESULT 15
LOCUS AR477463
DEFINITION Sequence 15 from patent US 6696292.
ACCESSION AR477463
PAT 14-MAY-2004
```

```

VERSION      AR477463.1  GI:47234921
KEYWORDS     Unknown.
SOURCE       Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 2067)
AUTHORS      Allen,S.M., Falco,S.C. and Thorpe,C.J.
TITLE        Genes encoding sulfo state assimilation proteins
JOURNAL      Patent: US 6696292-A 15 24-FEB-2004;
FEATURES     Location/Qualifiers
             1..2067
             source          /organism="unknown"
                             /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.:      2,66e-142      Length:      2067
Score:          1901.00        Matches:     359
Percent Similarity: 76.48%     Conservative: 106
Best Local Similarity: 59.05%  Mismatches:   143
Query Match:    54.71%        Indels:       0
DB:             6              Gaps:        0

US-10-762-049-18 (1-680) x AR477463 (1-2067)

QY      48  ProProGlnProPhePheLysSerLeuLysTySerLeuLysGluThrPhePheProAsp 67
      DB      8  CCACACCAGACCACTCCACAACTTAGGCACAGAGTCTCCGAAATCTTCTCCAGAT 67
QY      68  AspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPheMetLeuGlyLeuGln 87
      DB      68  GACCTCTCCACCGCTTTCAAGAACCACTCGCTTTAAAGATTCCTCTCGCAGCTTCAG 127
QY      88  PhePhePheProIlePheGluTrpAlaProLysTyThrPheGlnPheLeuLysAlaAsp 107
      DB      128  TATCTCTTCCCAATTTTCGACTGGGCCCACTACAACTTACCCTTCTCCGCTCTGAC 187
QY      108  LeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyAlaLys 127
      DB      188  CTCATCTCTGGCTCACCATTGCGAGCTCGCATTCCTCAGGGAATCAGTTATGCGAAG 247
QY      128  LeuAlaAsnLeuProIleLeuGlyLeuTyThrSerPheIleProProLeuIleTyThr 147
      DB      248  CTTGCCAACTTGGCACCATTCTTGGATATATTCGAGTTTGTCTCCCAATGATATAC 307
QY      148  AlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeu 167
      DB      308  TCGCTGCTGGAAGTCTAGACATCTTGGTGGTGGACCTGTTTCCATTGGCTCTTTGGTC 367
QY      168  MetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyThrLeuHis 187
      DB      368  ATGGGATCAATGTTAAGTGATAAAATTTCTTACACTCAAGAACCTATTCTCTATCTGGGA 427
QY      188  LeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArg 207
      DB      428  TTGGCTTTTACCCGCCATTCTTCTTGGTGGGTATTCGAAGCTTCTCTGGGTATATTAAG 487
QY      208  LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAla 227
      DB      488  CTAGGCTTCGTATTGATTCTTCTGCGAAGCAACGCTGGTGGATTCCAGCGCGGTGCT 547
QY      228  AlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHis 247
      DB      548  GCCATTATTGTGCTCACTGCGACGAGCTGAAGAGTTTACTTGAATAGTGCACTTTTACCAGC 607
QY      248  GlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrp 267
      DB      608  AAGATGCAAAATAATTCAGTAACGATCTCTGTGTTTCAAGCAAAAGACACGAGTGGTCATGG 667
QY      268  GluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyThrPheSer 287
      DB      668  CAACCACTTCTTTGGGATTCGGCTTCTCTTGGTGTGACCAACAGGCGACATTAGT 727

288  LysLysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeu 307
      DB      728  TTGAGAAACCAAACTATTCTGGGTTTCAGCAGCTGCCCATTTGACATCAGTATTCTG 787
QY      308  GlySerLeuLeuValTyThrHisAlaGluLysHisGlyValGluValIleGlyGlu 327
      DB      788  TCNACCACTTTAGTCTTCTTCTGAGAAATAGACTCATCAATTCAGTTATTTGGGCAC 847
QY      328  LeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyMet 347
      DB      848  TTACCAAGAGGAGTTAATCCACCATCAGCAACATGTTATCTTCAATGGTCCTTACTTG 907
QY      348  ThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAla 367
      DB      908  GGTCTTGTCTATCAAACTGGCATCATCAGGGATCTTATCTCTACTGAAGAAATGCA 967
QY      368  ValGlyArgSerPheAlaMetTyLysAsnTyAsnIleAspGlyAsnLysGluMetIle 387
      DB      968  GTAGGGAGAACATTTGCTTCACTTAAGAACTACCAGGTGGATGGAACAAAGAAATGATG 1027
QY      388  AlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyThrLeuThrThrGlyPro 407
      DB      1028  GCCATTGGTCTAATGAACATAGCTGGCTCGTGTCTTCTCATGTTATGTTACACGGGATCC 1087
QY      408  PheSerArgSerAlaValAsnTyAsnAlaGlyCysLysThrAlaAlaSerAsnIleIle 427
      DB      1088  TTTTCTCGATCGCTGTTAACTATATATGCTGAGCAGACACACAGTTTCAATATAATC 1147
QY      428  MetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyThrPro 447
      DB      1148  ATGGCTGAGCTGTCTAGTGACATCTTGTTCTCTCATGCTCTTCTTACTATACACCA 1207
QY      448  LeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyGluAla 467
      DB      1208  AATGTTGTCTTAGCGCCATTTATCATCTGCTGTGATTGGTCTAATAGATTATCAATCT 1267
QY      468  AlaIleHisLeuPheLysValAspLysPheAppPheValValCysMetSerAlaTyIle 487
      DB      1268  GCATATAAATTGGGAAGGTTGACAACTTGATTTCTTGGCCCTGTTTGGCTCTTTTTT 1327
QY      488  GlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeu 507
      DB      1328  GGGTCTCTTCTTATTTCACTGCTTGTAGTCTTGGTATAGCGGTTATCATATCAGTCTTC 1387
QY      508  ArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSer 527
      DB      1388  AAGATCTCTCTCATGCTCACTCGACCAACACTTTTGGTTTGGGGAATATACCAGGAACA 1447
QY      528  ValIleTyArgAsnValGluHisTyThrGlnAsnAlaLysHisValProGlyMetLeuIle 547
      DB      1448  CAATATTCCACACATAAACCAACCAATAACAAAAGCTTTAAGAGATTCCTTCTTCTCAT 1507
QY      548  LeuGluIleAspAlaProIleTyPheAlaAsnAlaSerTyThrLeuArgGluArgIleThr 567
      DB      1508  TTGGCTGTGTAGTCTCCAATCTATTTTGTCTAATCACTTATCTTCAAGAAAGGATCTG 1567
QY      568  ArgTyTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyThr 587
      DB      1568  AGATGGTTCGAGAGAGAGAGAGAGCATATAAAGCTATAATAGGAGCTCCATTGGAAGTGC 1627
QY      588  ValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607
      DB      1628  ATAATTTTAGACATGACAGCTGTCAGCCACAGACACAAAGTGGGCTTGCACACTTTATGT 1687
QY      608  GluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSer 627
      DB      1688  GAACCTTAGAAAGATGCTGGAGAGAGATCATTGAGTTTGTGCTGGCAATCTCTGTGGA 1747
QY      628  GluValMetLysLysLeuAsnLysSerTyPheGlnAsnHisLeuGlyLysTyTrpIle 647
      DB      1748  AATGTGATGAAAAAATTCGATAAGTCAACAACTTTGGATCTCTTTTGGATTAAAAAGGATC 1807
QY      648  TyrLeuThrValGluGluAlaVal 655
```

Db 1808 TATCTCAGTGGGAGAGCTGTG 1831
|||||

Search completed: August 31, 2005, 14:27:01
Job time : 8217 secs